DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY
Office for Scholarly Communication

Three Essays on Credit Risk Models and Their Bayesian Estimation

Citation

Kwon, Tae Yeon. 2012. Three Essays on Credit Risk Models and Their Bayesian Estimation. Doctoral dissertation, Harvard University.

Permanent link

http://nrs.harvard.edu/urn-3:HUL.InstRepos:9288549

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story .

Accessibility

©*2012 - Tae Yeon Kwon*

All rights reserved.

Dissertation Advisors: Professor Stephen Blyth Tae Yeon Kwon

Three Essays on Credit Risk Models and Their Bayesian Estimation

ABSTRACT

This dissertation consists of three essays on credit risk models and their Bayesian estimation. In each essay, defaults or default correlation models are built under one of two main streams in credit risk model study: the structural and the intensity models. The first essay studies the usefulness and methods to combine multiple securities information in a single firm asset process and to estimate its parameters under the structural model. The second essay investigates multi-firm correlated defaults, with special focus on industry-specific correlation under the intensity model. The third essay studies the use of multiple securities information to estimate the multi-firm correlated defaults model under both structural and intensity models.

In our first essay, sequential estimation on hidden asset value and model parameters estimation are implemented under the Black-Cox model. To capture short-term autocorrelation in the stock market, we assume that market noise follows a mean reverting process. For estimation, two Bayesian methods are applied in this essay: the particle filter algorithm for sequential estimation of asset value and the generalized Gibbs sampling method for model parameters estimation. The first simulation study shows that sequential hidden asset value estimation using option price and equity price is more efficient and accurate than estimation using only equity price. The second simulation study shows that by applying the generalized Gibbs sampling method, model parameters can be successfully estimated under the model setting that there is no closed-form solution. In an empirical analysis using eight companies, half of which are DowJones30 companies and the other half non-Dow Jones 30 companies, the stock market noise for the firms with more liquid stock is estimated as having smaller volatility in market noise processes.

In our second essay, the *frailty* idea described in Duffie, Eckner, Horel, and Saita (2009) is expanded to industry-specific terms. The MCEM algorithm is used to estimate parameters and

random effect processes under the condition of unknown hidden paths and analytically-difficult likelihood functions. The estimate used in the study are based on U.S. public firms between 1990 and 2008. By introducing industry-specific hidden factors and assuming that they are random effects, a comparison is made of the relative scale of within- and between-industries correlations. A comparison study is also developed among a without-hidden-factor model, a common-hidden-factor model, and our industry-specific common-factor model. The empirical results show that an industry-specific common factor is necessary for adjusting over- or under-estimation of default probabilities and over- or under-estimation of observed common factor effects.

Our third essay combines and extends works of the first two essays by proposing a common model frame for both structural and intensity credit risk models. The common model frame combines the merits of several default correlation studies which are independently developed under each model setting. Following the work of Duffie, Eckner, Horel, and Saita (2009), we apply not only observed common factors, but also un-observed hidden factor to explain the correlated defaults. Bayesian techniques are used for estimation and generalized Gibbs sampling and Metropolis-Hasting (MH) algorithms are developed. More than a simple combination of two model approaches (structural and intensity models), we relax the assumptions of equal factor effect across entire firms in previous studies, instead adopting a random coefficients model. Also, a novelty of the approach lies in the fact that CDS and equity prices are used together for estimation. A simulation study shows that the posterior convergence is improved by adding CDS prices in estimation. Empirical results based on daily data of 125 companies comprising CDS.NA.IG13 in 2009 supports the necessity of such relaxations of assumption in previous studies. In order to demonstrate potential practical applications of the proposed framework, we derive the posterior distribution of CDX tranche prices. Our correlated structural model is successfully able to predict all the CDX tranche prices, but our correlated intensity model results suggests the need for further modification of the model.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I am deeply grateful for the guidance of my thesis committee: Yoonjung Lee, Stephen Blyth, Tirthankar Dasgupta and Xiao-Li Meng. I acknowledge helpful comments from Steven Finch and Samuel Kou.

Most importantly, I thank my husband, my daughter, and my parents for their love and encouragement. I could not have written this without their support. This dissertation is dedicated to them.

0. BACKGROUND AND MOTIVATION

In the financial market, numerous and various market prices of financial securities are observed. Given these market prices, many investors and financial analysts ask a critical question, namely, what are market perceptions in the context of default probabilities? To answer this, research in the area of credit risk modeling has evolved by building a better statistical model describing defaults and trying to deduce prices from these models. The research in this dissertation is one such effort.

The literature in the credit risk area attempts to describe the default processes of debt as being primarily based on one of two types of models: the structural model or the intensity model. In this dissertation, we also expand models under one of these two types; therefore the general background and motivation for our three essays will first be introduced.

Structural models price corporate debts and equity as contingent claims on the underlying asset value of a firm. When the asset value of the firm falls below a certain threshold, the firm fails to meet its obligations to the debt holders, thus triggering a default event. The Merton model (Merton (1974)) provides a fundamental framework for various other structural models. The main idea behind Merton's model is to consider equity as a European-call-option-type contingent claim on asset. Black and Cox (Black and Cox (1976)) base their model on Merton's model, but they relax their assumption of default time.

Unlike structural models, intensity models specify neither firm value processes nor default boundaries explicitly. Defaults are instead modeled as stochastic events whose arrival rates are governed by the given intensities. The parameters governing intensities are typically assumed to depend on a set of market data.

Regardless of which model is being used, the starting point of the credit risk model is a given

market price. The prices observed in the financial market are numerous and various due to the large number of firms issuing multiple securities. Studies in this dissertation develop from single firm to multi-firm and from single security to multiple securities. Our study on default modeling starts from the multiple securities issued by a single firm which is described in Chapter 1. We then extend to multi-firms defaults analysis using the information of one security in Chapter 2. Finally, in Chapter 3, we combine all works in previous two research and then analyze multi-firm defaults using multiple securities.

The first essay conducted here is a cross-asset class research project with the consideration of each market noise based on the Black-Cox structural model. Structural models, in general, have potential advantages over intensity models in cross-asset class research. Conceptually, the capital structure of a firm can be evaluated as a whole in one consistent framework by structural models. Subsequently, methodologies for estimating models in a structural setting may deliver enhanced performances, since the relevant information can be pooled from the financial markets for other asset classes.

However, the default study based on a single-firm always has limitations in terms of default correlation. In the simplest case, default correlation is caused if one firm is a creditor of another. But primarily, and more generally, it is because the health of individual companies are linked together via industry-specific and/or general economic conditions.(see Lucas (1995), Zhou (2001)) The historical data also shows that all companies suffer or prosper together. For example, the pattern of yearly default rate for all U.S. corporate debt since 1900 shows the high concentration of defaults around 1914 and 1933. Many firms defaulted in those two depressions. All businesses tend to be adversely affected at the same time, because of their sensitivity to the general economy. (Lucas (1995))

Conversely, the default of one company can have an impact on other companies or general economic conditions. For example, during 2001, the year when Enron defaulted, the NASDAQ index fell 74% from its high and 171 large corporations declared bankruptcy, which was two times more than what took place in 2000. Bankruptcies were common throughout 2002, continuing into July, 2002 when WorldCom, the countrys second largest long-distance telecom company,

filed for bankruptcy(Smith and Walter (2006)). During the financial crisis that started in 2007, many financial institutions collapsed and finally in September, 2008 Lehman Brothers declared bankruptcy. However a more serious problem is that the impact of the bankruptcy of Lehman Brothers is not just constrained in the U.S. The fall of Lehman Brothers resulted in a 3.2% dip in the U.S. GDP within a span of six months, between the third quarter of 2008 and first quarter of 2009. The effect of the Lehman Brothers default in Italy, EU-27, Germany, and U.K. was even worse; these countries experienced a 4.4%, 4.1%, 5.8%, and 3.5% drop, respectively, during this period (see Daveri (2009)). Based on the fact that the U.S. GDP growth rate during the first six months in 2008 was only +1.14%, these rate drops were considered quite serious. This economic downturn brought by one firm's default might again affect the default risk of other firms.

From our second essay, we steer our study to multi-firm defaults analysis from the single firm default analysis. The key issues of the multi-firm defaults study is to find a way to capture default correlations more accurately. In practice, understanding default correlation correctly is important for understanding the distribution of a loan portfolio's loss and risk management. Particularly, in markets such as synthetic collateralized debt obligations (CDO), which are heavily reliant on the market for default swaps and often have basket structures, a modeling default correlation is necessary for pricing and rating. For example, the higher the correlation, the smaller the gap between the estimated risk of a CDO's AAA tranche and its equity tranche.

As our first multi-firm study, we build a model for correlated defaults under the intensity model in our second essay. The reason for switching to the intensity model from the structural model is because under the intensity model default correlation can be measured more easily by assuming commonly shared factors. As Duffie and Garleanu (2001), the intensity model assume that the default intensities of all firms are conditionally independent given the path of the state. The intensity model evolves by relaxing this conditional independent assumption. Collin-Dufresne and Helwege (2003) relax the conditional independence assumption, allowing default intensities to depend also on a common unobserved factor. In their model, a market-wise response to a credit event is due to investors' updating their beliefs about the unobserved factor after encountering a credit event. More recently, Duffie, Eckner, Horel, and Saita (2009) implement an intensity model

with a common unobserved factor named *frailty*. In our second essay, we expand this *frailty* idea to industry-specific terms and build the model under more relaxed assumptions allowing all different within-and between-industries correlations.

Finally, in our third essay, we extend previous studies in both directions: the multiple securities and the multiple firms. In terms of the multiple securities study, a typical structural model has an advantage. However, a typical intensity model has a merit in dealing with multi-firms. Therefore, we propose multi-firms defaults model frames and an estimation method using the multiple securities information under both intensity and structural credit risk model within the parallel model frame. We combine advantages and ideas in previous studies of intensity and structural correlated defaults models, then build one common structure for both. As a combining work of two different credit risk model studies, the main contribution of our third essay is that by proposing a common model structure for both structural and intensity models, we are able to compare the two under fair conditions. The relative importance of hidden common factors is compared to observed common factors in both models. Furthermore, a comparison is made of the performance of the CDX tranche price prediction between the two models.

However, as a first step in combining the work of two different credit risk model studies, we need to simplify some important points already suggested in our previous two essays. Even though we simplify the model with a common hidden factor instead of the industry-specific factors (proposed in the second essay), and we assume independent normal trading noise in the stock and CDS markets instead of mean-reverting process (proposed in the first essay), our third essay can be interpreted as synthesizing the work in our first two essays. We will leave relaxing such constraints as our future works.

Detailed statistical estimation methodologies adopted in each essay is different, but they all belong to the category of Bayesian analysis. The Bayesian approach is adaptable to cases where a closed-form solution is unavailable. Prediction is the other reason for using the Bayesian approach. By adopting the Bayesian method, we are able to extract more information on the hidden asset or intensity value itself as well as various financial securities. Because the Bayesian approach produces the posterior distribution of unknown parameters instead of one single value of

estimates, the distribution of a future observation known as a posterior predictive distribution can be easily derived. Furthermore, not only are the distribution of hidden asset or intensity values and prices used for estimation, but also the distribution of other securities not used in the estimation procedure, are all available.

1. STRUCTURAL CREDIT RISK MODEL WHEN MARKET PRICES ARE CONTAMINATED WITH NOISE

1.1 Introduction

Since Black and Scholes (1973) and Merton (1974) first considered equity as a call option on the firm's asset value, Merton's structural model has remained as the basic reference in pricing defaultable bonds. The major difficulty of the structural approach is that the firm's asset value is not observed directly. Through the observable equity prices, we are able to infer about the hidden asset values. However, since observed equity value is often contaminated by trading noise, arriving at inference becomes complicated. Faced with these challenges, this essay presents two types of estimation procedures of the Black and Cox model within a Bayesian framework. In the sense of cross-asset class research, sequential estimation of hidden asset value process is conducted first under the condition that all parameters are given. We then propose the method for estimation of all model parameters and asset processes together.

The contributions of this essay can be summarized as follows: First, we allow for trading noises in the observed equity prices, which follow a mean-reverting process. In markets where the trading noise effect exists, it will be ill-advised to ignore its presence. In recent years, research on trading noise has provided reasons for this fact. For instance, Ait-Sahalia and Zhang (2005a) analyze the effect of the trading noise on how frequently one should sample the equity price. Ait-Sahalia and Zhang (2005b) and Bandi and Russell (2006) show the effects of trading noise on volatility estimation. Even though it is a well known fact that observed equity prices can diverge from their equilibrium values due to several reasons (market illiquidity and model mis-specification), because of the complex estimation, trading noise in the stock market has not been fully modeled together with asset prices prior to Duan and Fulop (2009). In this essay, we extend

Duan and Fulop (2009)'s approach by assuming that trading noise in the equity market follows a mean reverting process instead of an independent normal distribution. The mean-reverting process assumption is closer to real stock market behavior, which shows short-term dependence rather than the independent normality.

Second, we estimate parameters (of asset and noise processes) and unknown asset processes together under the Black-Cox structural model (Black and Cox (1976)). Previous studies on simultaneous estimation of asset process and its parameters were based on the Merton model which gives a relatively simple inverse function of equity and asset values. The iterative scheme (Vassalou and Xing (2004)) and the well-known KMV method (or transformed-data MLE method described in Duan (2000), Duan (1994)) are all based on the Merton model. However, the Black-Cox model has a more relaxed assumption on default time than the Merton model. The main contribution of the Black-Cox model is to allow a default to occur anytime prior to the maturity of the bond. In practice, under the usual bond contract, a firm needs to pay annual interest to the bond holder, usually in semi-annual tranches. Default thus can occur anytime before the bond matures, when the firm fails to pay this coupon payment.

Third, we model the prices of multiple asset classes (debts, equity, and options on equity), in order to gain a better understanding of a firm's capital structure as a whole and also to enhance the precision of the estimation. In the financial market, corporations tend to issue multiple classes of securities to raise their capital. For instance, equity, corporate bonds, and various types of hybrid securities comprise significant financial markets. Therefore, the stock market alone provides only a limited view of the firm's capital structure. To improve our estimation, we adopt the cross-asset class approach by adding option price information into our model.

The cross-asset class research already has gained attention in other fields of financial research during recent years. For instance, Ni, Pan, and Poteshman (2008) illustrate that option prices tend to be more informative than stock prices in estimating stochastic volatility models. In addition, they show that trading volume in the option markets generally contains more information about future realized volatilities than trading volume in the stock markets. Using the particle filtering algorithm, Johannes, Polson, and Stroud (2007) show that volatility estimation with both option

price and equity price is more efficient and accurate than estimation using only equity price. Hull, Nelken, and White (2004) propose a novel attempt in the direction of cross-asset class analysis, estimating the parameters that govern default risks via implied volatility curves observed in the option markets. In this sense, we adopt the cross-asset class approach to credit risk modeling. By pooling information across different markets, the cross-asset class approach offers a more precise picture to investors of a firm's asset structure

By applying the Bayesian method, all unknown asset processes, their parameters and noise process parameters are successfully estimated. Under the existence of noise and noise model parameters, it is impossible to adopt an iterative method as has been applied in previous studies. Furthermore, the option pricing solution does not have a closed form solution. We propose the Bayesian estimation method, which can be applied in these complicated model setups. The main methodology applied in sequential estimation of the hidden asset process is the particle filter (also known as the bootstrap filter) (Gordon, Salmond, and Smith 1993). This method applies the concept of sampling-important-resampling (SIR) (Rubin 1987). In our model, multiple classes of assets are linked through the unobserved underlying firm value process. By running the filtering algorithm, the conditional distribution of the underlying asset value is approximated and recursively updated, given observed equity and option prices. For model parameter estimations, the generalized Gibbs sampling algorithm (Liu and Sabatti 2000) is adopted. The main idea of the generalized Gibbs sampling method is in line with simple Gibbs sampling, but it introduces an update that changes correlated parameters simultaneously. When parameters are highly correlated, the sampler should be able to make them converge faster by moving them together.

The rest of this essay is constructed as follows. Section 1.2 introduces the model we adopt and details on how to price debts, equity, and options on equity in the Black-Cox model setting with an added layer of the market microstructure noise. Section 1.3 provides the particle filtering algorithm for a sequential hidden asset process estimation. Section 1.4 provides a generalized Gibbs sampling method for parameter estimation. Section 1.5 shows the results of parameter estimation with real data. Finally, the essay ends with discussions for further studies in Section 1.6.

1.2 The Model

The model we adopt in this essay can be summarized as follows. First, we adopt the Black-Cox model for pricing debt and equity. Second, in pricing them, we consider the market noise following the mean reverting process to capture short-term autocorrelation exhibited in the stock market.[1] Finally, to enhance estimation performance, we add the option on equity market information into our model.

1.2.1 Basic Model: Equity Prices in the Presence of Market Noise

Under the structural credit risk model, defaults events are determined by the asset price. When the asset value of the firm falls below a certain threshold, the firm fails to meet its obligations to the debt holders, thus triggering a default event. Therefore, in order to analyze and predict the default behavior, correct estimation of asset process is the most important issue in the structural credit risk model.

As in a typical structural credit risk model, let us consider a firm with its value of the asset V_t following a geometric Brownian motion under the probability measure P:

$$dV_t = \mu V_t dt + \sigma_V V_t dW_t^{PV}, \tag{1.1}$$

The processes W^{PV} is standard Brownian motion. μ is the mean parameter and the volatility parameter σ_V is some positive constant to be estimated.

However, the default analysis under the structural model is not a simple drift μ_V and volatility σ_V estimation problem. First, the fact that we do not observe the asset value V_t of the firm complicates the default analysis. In order to deal with this complication, several structural credit risk models have derived the equity pricing formulas as a function of asset and parameters governing asset value dynamic. Based on this derivation, an iterative scheme (Vassalou and Xing (2004))

[1] We will keep using *stock* and *equity* without any distinction. They are different terminologies in finance, but we calculate equity as a product of stock price and number of outstanding. Equity referred in this essay is actually stock market information.

then gives unobserved asset process using the observed equity prices. In this essay, we also use the equity price to derive the asset value, but we take into consideration another complexity in real market as follows.

Second complexity is due to market noise. Any of well-derived models cannot perfectly explain reality. There may be gaps between observed equity values and model-derived equity values. The market noise literature indeed strongly suggests that noise should be expected (Ait-Sahalia and Zhang (2005a), Ait-Sahalia and Zhang (2005b), Bandi and Russell (2006), and Duan and Fulop (2009)). In this sense, market noise is incorporated into our model to adjust this model mis-specification error. It also reflects short-term discrepancies in the supply and demand in the stock market and bid-ask prices.

In order to incorporate market noise into the model, we assume multiplicative error structure. The log equity price $\ln S_t$ is

$$\ln S_t = \ln S_{Model,t}(V_t, t, \Theta_V) + Z_t, \tag{1.2}$$

where $S_{Model,t}(V_t, t, \Theta_V)$ is model-derived (theoretical) equity value and Z_t is market noise. The model-derived equity values $S_{Model,t}$ is a function of time t, asset value V_t, and the parameter governing asset dynamic $\Theta_V = (\mu, \sigma_V)$. We will discuss each $S_{Model,t}$ and Z_t term in equation (1.2) with more details as follows.

Z_t: Mean Reverting Market Noise Process

The simplest way to model the error is to assume an independently and normally distributed error, as in Duan and Fulop (2009). However, the stock market is unique in that stock prices (or return) tend to exhibit short-term auto-correlations (for example, Lo and MacKinlay (1988), Fama and French (1988), Poterba and Summers (1988), Conrad and Kaul (1989), Jegadeesh (1990), Lehmann (1990), and Gaunt and Gray (2003)). To make our model more consistent with empirical observations, we model market noise with a mean-reverting process as follows:

$$dZ_t = -\theta_Z Z_t dt + \sigma_Z dW_t^{PZ}, \tag{1.3}$$

where θ_Z and σ_Z are positive (unknown) constants. W_t^{PZ} is a standard Brownian motion under the probability measure P and it is assumed to be independent with dW^{PV} in equation (1.1).

$\ln S_{Model,t}(V_t, t, \Theta_V)$: *Equity Prices under the Black-Cox Model*

The other term comprising the equation (1.2) is the theoretical equity value which is $\ln S_{Model,t}(V_t, t, \Theta_V)$. The Black-Cox model is adopted in this essay. Past research on unknown asset process estimation has assumed the Merton model based on the relationship between observed equity and unobserved asset values represented by the simple inverse function of the Black-Scholes-Merton formula. However, in practice, defaults do not always occur at the maturity of the bond. Whenever a firm fails to pay the interest or dividend, default can always happen. Moreover, according to Moody's Default Risk Service, there are various reasons for defaults: bankruptcy, distressed exchange, dividend omission, grace-period default, indenture modification, missed interest payment, missed principal and interest payments, missed principal payment, payment moratorium, and suspension of payments. The Black-Cox model is applied in order to relax an assumption on default time in spite of the complication it may present in mathematical derivation and application.

Before we move on to the next step of the model, we will introduce some basics in pricing debt and equity under the Black-Cox model settings. V_t is assumed to be following a geometric Brownian motion under the probability measure P as in equation (1.1). Under the risk neutral measure (or the equivalent martingale measure) Q [2], it follows

$$dV_t = rV_t dt + \sigma_V V_t dW_t^{QV}, \tag{1.4}$$

where r is a known constant risk-less rate. The processes W^{QV} are standard Brownian motions under Q.

Now assume that the firm at time 0 has issued two types of claims: debt and equity. Debt is

[2] Risk-neutral measures make it easy to express the value of a derivative in a formula. In mathematical finance, a risk-neutral measure is a prototypical case of an equivalent martingale measure. It is heavily used in the pricing of financial derivatives due to the fundamental theorem of asset pricing, which implies that in a complete market a derivative's price is the discounted expected value of the future payoff under the unique risk-neutral measure.

zero-coupon bond with a face value D and maturity date is T_D. The default boundary L_t is given as follows:

$$L_t = L_0 \exp\left(-\gamma(T_D - t)\right), \tag{1.5}$$

for some positive constants L_0 and γ. When asset value falls below a certain boundary, the firm defaults thus the default time τ is defined by

$$\tau = \inf\{t \geq 0 : V_t \leq L_t\}. \tag{1.6}$$

Debt has a claim priority over equity and equity holders are protected by limited liability. If $\tau > T_D$, in other words, no default has occurred by the time a bond matures then the bond holder pays the face value of the debt. If assets at time T_D are worth less than the face value of debt D, the bond holder then takes over the remaining asset. The payoff to the bond holder at maturity date T_D and when asset value at T_D is V_{T_D}, which is referred to by $B_{BC}(V_{T_D}, T_D)$, and to an equity holder, $S_{BC}(V_{T_D}, T_D)$, are given by:

$$\begin{aligned} B_{BC}(V_{T_D}, T_D) &= \min\left(D, V_{T_D}\right) \mathbf{1}_{\{\tau > T_D\}}, & (1.7) \\ S_{BC}(V_{T_D}, T_D) &= \max\left(V_{T_D} - D, 0\right) \mathbf{1}_{\{\tau > T_D\}}. & (1.8) \end{aligned}$$

If the default occurs ($\tau \leq T_D$), then the bond holder takes over the firm. The payoff to the bond holder at default date τ and when asset value at τ is V_τ, which is referred to by $B_{BC}(V_\tau, \tau)$, and to an equity holder, $S_{BC}(V_\tau, \tau)$, are given by:

$$\begin{aligned} B_{BC}(V_\tau, \tau) &= L_\tau \mathbf{1}_{\{\tau \leq T_D\}}, & (1.9) \\ S_{BC}(V_\tau, \tau) &= 0 \mathbf{1}_{\{\tau \leq T_D\}}. & (1.10) \end{aligned}$$

Given these pay-off functions, Lando (2004) carries out the calculation and derives a closed form solution for $B_{BC}(V_t, t; T_D, D)$, which is the estimated price of the defaultable bond at time t. Detailed solution and derivation are given in Appendix A.1. Under the Black-Cox model, we

define $\ln S_{Model,t}(V_t,t,\Theta_V)$ in equation (1.2) as $g_S(V_t,t,\Theta_V)$, i.e.

$$\ln S_{Model,t}(V_t,t,\Theta_V) \equiv g_S(V_t,t,\Theta_V) = \ln\left(V_t - B_{BC}(V_t,t;T_D,D)\right) \tag{1.11}$$

In comparison with the Merton model, the Black-Cox model leads to higher bond prices and lower spreads, which is consistent with the boundary representing a safety covenant. From the equity point of view, the equity owners in the Merton model have a European call option; however, in the Black-Cox model they have a down-and-out call option.

Now we are ready to use the equity price process in estimation by setting the link function between observed equity prices and unknown asset values. Another new aspect of our model is that option prices information is additionally used for estimation as a cross-asset class research. In the next subsection, in preparation for using the option prices information in estimation, we will derive the theoretical price of option on equity when equity price is contaminated with market noise as defined earlier.

1.2.2 Multiple Asset Class : Pricing Options on Equity

Generally, asset value information is not open to the public. In order to make a better investment decision, investors need to infer the firm's asset status based on published markets information. So far, we set our model to use equity prices in asset value inference. However, we are able to enhance the precision of the estimation if information from more varied asset classes issued by a firm is available. Therefore, an understanding of a firm's capital structure as a whole, based on information from several financial markets, is an important step. In this essay, as additional market information, we use the option on equity market. In order to derive a price of option on equity in the presence of market noise, we perform following calculation.

Let us consider a European call option on the underlying equity with a maturity $T_C \leq T_D$ and a strike price K. We assume that observed option prices C_t are exposed to measurement errors ε_t.

$$C_t = g_C(S_t,V_t,t,\Theta_V,\Theta_Z) + \varepsilon_t, \tag{1.12}$$

where $\Theta_Z = (\theta_Z, \sigma_Z)$. $g_C(S_t, V_t, t, \Theta_V, \Theta_Z)$ is the theoretical option price derived from the model. $\varepsilon_t \sim N(0, \sigma_\varepsilon)$ independently. $g_C(S_t, V_t, t, \Theta_V, \Theta_Z)$ is computed by

$$g_C(S_t, V_t, t, \Theta_V, \Theta_Z) = E^Q \left[\exp(-r(T_C - t)) \max(S_{T_C} - K, 0) \right], \tag{1.13}$$

where E^Q is the expectation under the Q-measure. In the presence of stock market noise, the simple Black-Scholes formula cannot be adopted to derive a closed form solution for option price. We thus approximate the theoretical option price $g_C(S_t, V_t, t, \Theta_V, \Theta_Z)$ via Monte Carlo simulation as in Hull and White (1987).

To obtain Q dynamics of S_t, we need to derive its P dynamics first. From the equations (1.2) and (1.11), the market price of equity becomes,

$$\ln S_t = g_S(V_t, t, \Theta_V) + Z_t. \tag{1.14}$$

Applying Itô's formula (Ito (1951)), we have

$$\begin{aligned} d\ln S_t &= dg_S(V_t, t, \Theta_V) + dZ_t \qquad (1.15) \\ &= \frac{\partial g_S}{\partial t}(V_t, t, \Theta_V)dt + \frac{\partial g_S}{\partial v}(V_t, t, \Theta_V)dV_t + \frac{1}{2}\frac{\partial^2 g_S}{\partial v^2}(V_t, t, \Theta_V)(dV_t)^2 \\ &\quad -\theta_Z Z_t dt + \sigma_z dW_t^{PZ} \\ &= \frac{\partial g_S}{\partial t}(V_t, t, \Theta_V)dt + \frac{\partial g_S}{\partial v}(V_t, t, \Theta_V)\left[\mu V_t dt + \sigma_V V_t dW_t^{PV}\right] + \frac{1}{2}\frac{\partial^2 g}{\partial v^2}(V_t, t, \Theta_V)\sigma_V^2 V_t^2 dt \\ &\quad -\theta_Z \left(\ln(S_t) - g_S(V_t, t, \Theta_V)\right) dt + \sigma_z dW_t^{PZ}, \end{aligned}$$

Since the discounted S_t is a martingale under Q and its volatility term remains unchanged under both measures P and Q, we now conclude that under Q, S_t follows

$$\begin{aligned} d\ln S_t &= \left(r - \frac{1}{2}\left(\frac{\partial g}{\partial v}(V_t, t, \Theta)\sigma_V V_t \right)^2 - \frac{1}{2}\sigma_Z^2 \right) dt \qquad (1.16) \\ &\quad + \frac{\partial g}{\partial v}(V_t, t, \Theta)\sigma_V V_t dW_t^{QV} + \sigma_z dW_t^{QZ}, \end{aligned}$$

where W_t^{QV} and W_t^{QZ} are independent Q-standard Brownian motions. By simulating S_t paths under Q dynamics, we can get expected value in equation (1.13).

In this section, we derived the theoretical value of equity and option under our model assumptions. We also introduced the distribution assumptions on noise factors in equity and option market respectively. Based on this derivation and distribution assumptions, we are able to complete likelihood functions of unknown quantities. The next two sections will address the Bayesian estimation procedures on the unknown quantities by utilizing the derived likelihood functions.

1.3 The Sequential Estimation on Hidden Asset Process

One of the main challenges in applying structural models to financial market data is the fact that the underlying asset value process is unobservable. Furthermore, at each time t, market values of equity and option are known only up to the time t, which means that the information needs to be updated sequentially. In this section, with known model parameters, we apply the particle filter algorithm, described in K.P. and Shephard (1999) and Johannes, Polson, and Stroud (2007) to update the information about the underlying asset value process recursively from the observed times series of equity and option prices. All of the estimation will now be under discretized time frame.

1.3.1 Estimation Procedure

To see how much option price information can improve the estimation performance, we first estimate the hidden asset value process only with equity prices. Details about the particle filter algorithm are given in Appendix A.2.

With known parameters $\Theta = \{\mu, \sigma_V, \theta_Z, \sigma_Z\}$, we observe the time series of stock prices $S = \{S_t; t = 1, ..., T\}$ and have the hidden asset process to be estimated $V = \{V_t; t = 1, ...T\}$. In addition, market noise process $Z = \{Z_t; t = 1, ...T\}$ is also unobservable.

In the process of following the particle filter algorithm steps, we confront the problem that re-sampling weight $w^j \approx f_{t+1}(S_{t+1}|V_{t+1}^{(*j)}, Z_{t+1}^{(*j)})$ is a δ-function, which takes only 0 or 1 values. In order to solve this problem, the algorithm is adjusted as follows:

$$\begin{array}{lccccccccc}
\text{Observed variables:} & S_1 & \rightarrow & S_2 & \rightarrow & ... & S_{T-1} & \rightarrow & S_T \\
 & \uparrow & & \uparrow & & ... & \uparrow & & \uparrow \\
\text{State variables:} & (V_1,Z_1) & \rightarrow & (V_2,Z_2) & \rightarrow & ... & (V_{T-1},Z_{T-1}) & \rightarrow & (V_T,Z_T)
\end{array}$$

1. Given $\{V_t^{(j)}\}_{i=1}^{M}$, draw $V_{t+1}^{(*j)}$ from $q_{t+1}(V_{t+1}|V_t^{(j)})$ where j=1,...M.

2. Set $Z_{t+1}^{(*j)}$ such that $f_{t+1}(S_{t+1}|Z_{t+1}^{(*j)},V_{t+1}^{(*j)}) > 0$.
   (i.e., find $Z_{t+1}^{(*j)}$, which makes $\ln(S(t+1)) = g(V_{t+1}^{(*j)},t,\Theta) + Z_{t+1}^{(*j)}$.)

3. Weight each draw by
$$w^{(j)} \propto |J| k_{t+1}(Z_{t+1}^{(*j)}|Z_t^j).$$

4. Resample from $\{(V_{t+1}^{(*1)},Z_{t+1}^{(*1)}),(V_{t+1}^{(*2)},Z_{t+1}^{(*2)}),..(V_{t+1}^{(*M)},Z_{t+1}^{(*M)})\}$ with probability proportional to $w^{(j)}$.

J is the Jacobian transformation of $Z_t = \ln(S_t) - g(V_t,t,\Theta)$. q_{t+1} is the density function of V_{t+1} given V_t and k_{t+1} is the density function of Z_{t+1} given Z_t, where

$$V_{t+1}|V_t^{(j)} \sim LogNormal\left(\ln(V_t^{(j)}) + (\mu - \frac{1}{2}\sigma_V^2)\Delta t, \sigma_V\sqrt{\Delta_t}\right), \tag{1.17}$$

$$Z_{t+1}|Z_t \sim Normal\left(\exp(-\theta_Z\Delta t)Z_t, \sigma_Z\sqrt{\frac{1}{2\theta_Z}(1-\exp(-2\theta_Z\Delta t))}\right). \tag{1.18}$$

After incorporating additional market information and option on equity prices, the re-sampling weight changes to

$$w^{(j)} \propto |J| k_{t+1}(Z_{t+1}^{(*j)}|Z_t^j) l_{t+1}(C_{t+1}|S_{t+1},V_{t+1}^{(*j)}), \tag{1.19}$$

because

$$f_{t+1}^{SC}(S_{t+1},C_{t+1}|V_{t+1},Z_{t+1}) = f_{t+1}(S_{t+1}|Z_{t+1},V_{t+1}) l_{t+1}(C_{t+1}|S_{t+1},V_{t+1}), \tag{1.20}$$

where C_t is a function of the observed price S_t and the state V_t.; l_{t+1} is the density function of C_{t+1} when $S_{1:t+1}$ and V_{t+1} are given, where

$$C_{t+1}|S_{t+1},V_{t+1} \sim Normal(g_S(S_{t+1},V_{t+1},t,\Theta_V,\Theta_Z),\sigma_\varepsilon) \tag{1.21}$$

1.3.2 Simulation-Based Result

In order to illustrate the effect of additional uses of option prices, we conduct a simulation study. First, we generate asset process V and market noise process Z with the parameters: The parameter values are chosen in a way that is consistent with real data. We use median values (rounded) obtained from real data (eight firms which will be described in section 1.5). We set $\mu = 0.1, \sigma_V = 0.1, \sigma_Z = 0.1$, and $\theta_Z = 30$. σ_ε is set to be 0.02; The time length of simulated data is 10 days and we use the Euler scheme; A starting value of asset V_1 is set to be 200, and a face value of bond is set to be 100 with two-years maturity. For the default boundary L_t, we set $L = 0.8V_1$ and $\gamma = 0.1$. The stock and option prices processes are then generated based on equation (1.2) and equation (1.12) respectively. Using the same parameters, this data generation procedure is repeated 105 times, providing 105 independent data sets. Finally, for each simulated data set, we draw 1000 posterior samples of hidden asset processes.

Then the path of error between the true hidden process and the posterior mean process, and the 95% posterior interval and posterior standard deviation paths for all 105 simulated data sets are calculated. With the mean of these paths on each day (time), the filter performance is summarized in two ways. First, the mean-squared error (MSE), which measures the mean-squared difference between true value and posterior mean, shows the precision of estimates. Second, the types of statistics are 95% posterior interval (PI) length and posterior standard deviation (SD) at each time $t = 2, ...10$. These two statistics are for measuring efficiency of estimates. Table 1.1 shows the average difference of MSE, length of 95% PI, and SD between two methods: estimation with stock prices information and estimation with stock and option prices. In order to show the statistical significance in the MSE, LPI, and SD decreases by adding option prices, paired t-tests are conducted with each statistics mean difference in each day (df=104) and also for all time period (df=105*9-1=944).

Negative MSE difference between MSE_{SC} and MSE_S indicates that estimation using both stock and option prices is more precise than using only stock prices. Table 1.1 also shows that, in general, posterior standard deviation and length of 95% prediction interval decrease by adding option price information. The mean difference of MSE for the entire time period (t=2,...10) is -0.034 and

Table 1.1: **Asset Process Estimation Results with the Simulated 105 Data Sets with Time Length=10, Time-by-Time.** Because 105 sample paths are simulated, we give the average of all estimation results: MSE, length of PI and SD. The MSE, length of PI and SD difference are calculated in each day. MSE_{SC} refers to average of all 105 MSE of estimation using stock and option prices information, MSE_S is average 105 MSE of estimation using stock prices information, LPI_{SC}, LPI_S are average length of 95% posterior interval using each method. SD_{SC}, SD_S are average of posterior standard deviation using each method. In order to show the statistical significance in the MSE, LPI, and SD decreases by adding option prices, paired t-tests are conducted with each statistics mean difference in each day (df=104) and also for all time period (df=105*9-1=944). The values in the parenthesis are one-sided paired t-test p-values. (* significant at 10%, ** significant at 5%.)

Time	$MSE_{SC} - MSE_S$	$LPI_{SC} - LPI_S$	$SD_{SC} - SD_S$
$t = 2$	-0.005(0.105)	+0.018(0.927)	+0.002 (0.726)
$t = 3$	-0.003(0.389)	-0.056(0.011)**	-0.014 (0.016)**
$t = 4$	-0.027(0.073)*	-0.043 (0.101)	-0.009 (0.079)*
$t = 5$	-0.030(0.138)	-0.059 (0.050)**	-0.006 (0.247)
$t = 6$	-0.051(0.082)*	+0.000 (0.502)	-0.001 (0.457)
$t = 7$	-0.055(0.062)*	-0.005 (0.455)	+0.004 (0.661)
$t = 8$	-0.051(0.135)	+0.003 (0.522)	+0.005 (0.667)
$t = 9$	-0.018(0.380)	-0.045 (0.192)	-0.001 (0.454)
$t = 10$	-0.066(0.118)	-0.065 (0.120)	-0.010 (0.186)
all t(from 2 to 10)	-0.034(0.003)**	-0.028 (0.021)**	-0.003 (0.126)

it is significantly less than 0 (p-value is 0.003, df=944). The decrease in length of posterior interval is also significant (p-value=0.021, df=944). This decrease shows that, by adding the option price, we are able to improve estimation efficiency. However, in each day, this improvement was not significant enough. In terms of MSE, any decrease is not significant at 5%.

Figure 1.1 and Table 1.2 are results from the long simulated sample path with a time length T=125 (6 months daily data). Instead of averaging over sample paths, we take the mean-squared error, posterior standard deviation, and 95% prediction interval length over the time period. Similar to previous results, by adding option price information, we gain precision and efficiency in estimation.

Both simulation results show that by using more information from different financial markets, we can improve the performance of hidden value estimation. The improvement, however, was marginal. When the model parameters are known, option price does not reveal substantially new information about the hidden asset value. However, it require extensive computing time to incor-

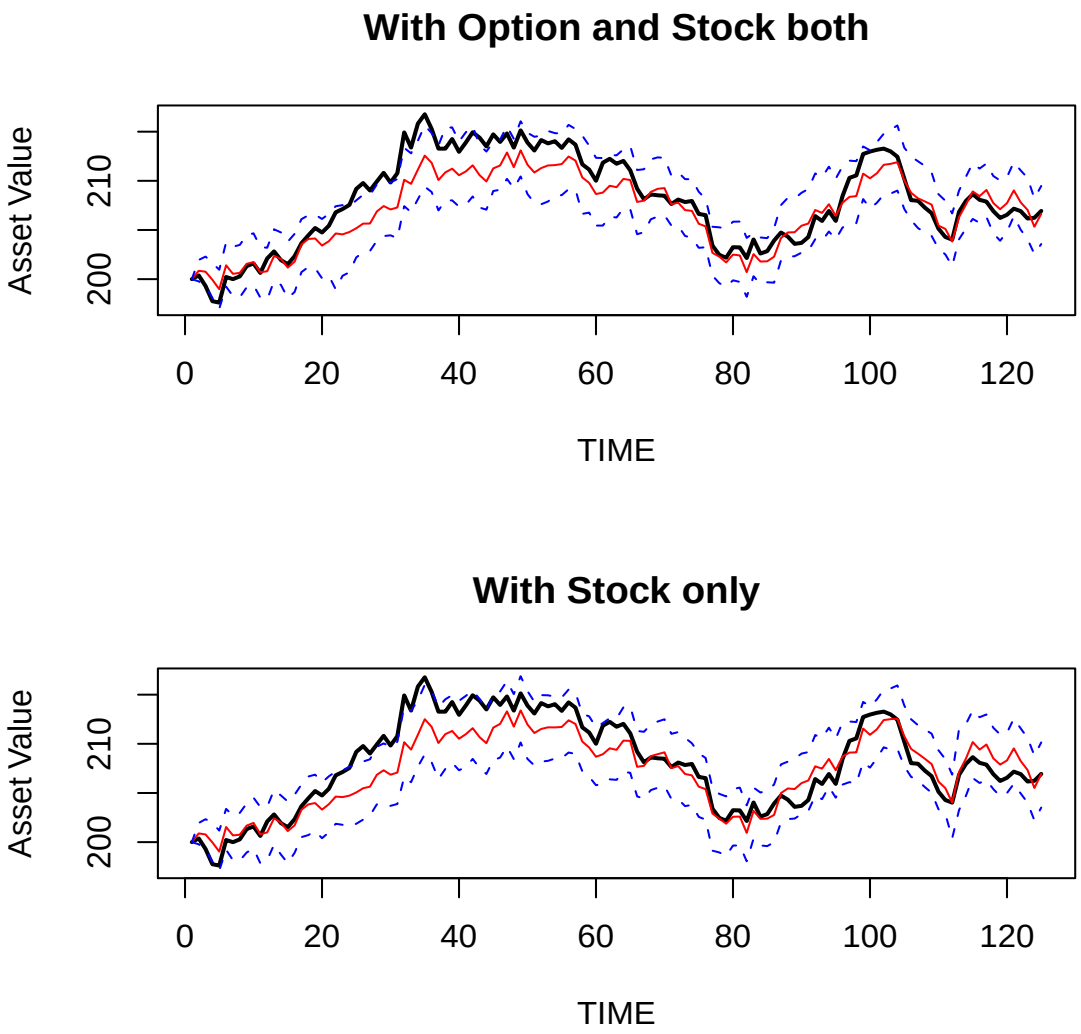

Figure 1.1: **Estimated Asset Process in the Simulated Data Set with Time Length=125.** The top panel is the estimated asset process using stock and option prices information and the bottom panel is the estimated asset process using only stock prices information. The solid *black* line is true value, the solid *red* line is posterior mean and the *blue* dotted lines are 95% PI.

Table 1.2: **Asset Process Estimation Results with the Simulated Data Set with Time Length=125.** We calculate the means of MSE , length of 95% PI, and posterior SD paths when we use both option and stock prices information and when we use only the stock prices information in estimation.

	Use Option, Stock both	Use Stock only
MSE	3.665957	3.74386
Mean of posterior SD	1.500945	1.564889
Mean Length of 95% PI	5.844097	6.093491

porate the option price into the model, because the closed-form solution for option price is not available. In the next section, in parameter estimation, only the stock value information is used because of the computational tractability.

1.4 Parameter Estimation

In the previous section, we assumed that model parameters are all known. However, such an assumption is unrealistic because the asset process itself is not observable. Moreover, under the condition of the existing market noise process, the iterative method, which was used earlier in the literature, cannot be adopted. We propose the Bayesian method to estimate the unknown asset process, its parameters, and the noise process parameters all together. In this section, the estimation procedure will be discussed in detail.

1.4.1 Simple Gibbs Sampling Method

First, we use a simple Gibbs sampling method to estimate parameter $\Theta = \{\mu, \sigma_V, \theta_Z, \sigma_Z\}$, and asset value process $V = \{V_t; t = 1, ...T\}$, using equity values $S = \{S_t; t = 1, ...T\}$.

To make the Markov Chain Monte Carlo sampling more efficient, we re-parameterize θ_Z by $\ln\theta_Z$ and V_t by $\ln V_t$. To implement the Gibbs sampling, two conditional distributions are necessary, $p_\Theta(\Theta|S, \ln V)$ and $p_{\ln V}(\ln V|S, \Theta)$:

$$p_\Theta(\Theta|S, \ln V) \propto \eta(\Theta)h(S, \ln V|\Theta), \tag{1.22}$$

$$p_{\ln V}(\ln V|S, \Theta) \propto h(\ln V, S|\Theta), \tag{1.23}$$

where $\eta(\Theta)$ is the prior distribution of parameters. In our model, there is one more unobserved process, namely, the market noise process $Z = \{Z_t; t = 1, ...T\}$; thus $h(\ln V, S|\Theta)$ is derived as follows:

$$\begin{aligned} h(\ln V, S|\Theta) &= \int h_Z(S, Z, \ln V|\Theta) dZ \qquad (1.24) \\ &= \int \prod_{t=1}^{n} q_t(\ln V_t|\ln V_{t-1}, \Theta) f_t(S_t|\ln V_t, Z_t, \Theta) k_t(Z_t|Z_{t-1}, \Theta) dZ \\ &= \prod_{t=1}^{n} \Phi\{\ln V_t; \mu_1(\ln V_{t-1}, \sigma_V, \mu), \sigma_1(\sigma_V, \mu)\} \times \\ &\qquad \Phi\{\ln S_t - g(V_t, t, \Theta); \mu_2(S_{t-1}, V_{t-1}, \sigma_V, \theta), \sigma_2(\theta, \sigma_Z)\}, \end{aligned}$$

where

$\Delta t = T/n$,

$\Phi\{x; \mu, \sigma\}$: pdf of x when $X \sim Normal(\mu, \sigma)$,

$\mu_1(\ln V_{t-1}, \sigma_V, \mu) = \ln V_{t-1} + (\mu - \sigma_V^2/2)\Delta t$,

$\sigma_1(\sigma_V, \mu) = \sigma_V \sqrt{\Delta t}$,

$\mu_2(S_{t-1}, V_{t-1}, \sigma_V, \theta_Z) = \exp(-\theta_Z \Delta t)(\ln S_{t-1} - g(V_{t-1}, t-1, \Theta))$,

$\sigma_2(\theta_Z, \sigma_Z) = \sigma_Z \sqrt{\frac{1-\exp(-2\theta_Z \Delta t)}{2\theta_Z}}$.

To obtain Monte Carlo samples from the joint posterior distribution described in equations (1.22) and (1.23), we iterate the following conditional sampling steps, starting from an initial configuration.

Step 1: Draw $\Theta = (\mu, \sigma_V, \ln \theta_Z, \sigma_Z)$ from $\eta(\Theta) h(\ln V, S|\Theta)$, so for each individual parameter, the target distribution is

- $\mu \sim [\mu|\sigma_V, \ln \theta_Z, \sigma_Z, \ln V, S] \propto \eta_\mu \prod_{t=1}^{n} \Phi\{\ln V_t; \mu_1(V_{t-1}, \sigma_V, \mu), \sigma_1(\sigma_V, \mu)\}$

- $\sigma_V \sim [\sigma_V|\mu, \ln\theta_Z, \sigma_Z, \ln V, S]$

$$\propto \quad \eta_{\sigma_V} \prod_{t=1}^{n} \Phi\{\ln V_t; \mu_1(V_{t-1}, \sigma_V, \mu), \sigma_1(\sigma_V, \mu)\} \times$$
$$\Phi\{\ln S_t - g(V_t, t, \Theta); \mu_2(S_{t-1}, V_{t-1}, \sigma_V, \theta_Z), \sigma_2(\theta_Z, \sigma_Z)\}$$

- $\ln\theta_Z \sim [\ln\theta_Z|\mu, \sigma_V, \sigma_Z, \ln V, S]$

$\propto \eta_{\ln\theta_Z} \prod_{t=1}^{n} \Phi\{\ln S_t - g(V_t, t, \Theta); \mu_2(S_{t-1}, V_{t-1}, \sigma_V, \theta_Z), \sigma_2(\theta_Z, \sigma_Z)\}$

- $\sigma_Z \sim [\sigma_Z|\mu, \sigma_V, \ln\theta_Z, \ln V, S]$

$\propto \eta_{\sigma_Z} \prod_{t=1}^{n} \Phi\{\ln S_t - g(V_t, t, \Theta); \mu_2(S_{t-1}, V_{t-1}, \sigma_V, \theta_Z), \sigma_2(\theta_Z, \sigma_Z)\}$

Step 2: Draw $\ln V = (\ln V_1, \ln V_2, .. \ln V_t, ... \ln V_n)$ from $p(\ln V_{1:n}|S_{1:n}, \Theta) \propto h(\ln V_{1:n}, S_{1:n}|\Theta)$, so for individual $\ln V_t$, the target distribution is

$$\ln V_t \sim [\ln V_t | \ln V_{t-1}, \ln V_{t+1}, S, \Theta]$$
$$\propto \quad \Phi\{\ln V_t; \mu_1(V_{t-1}, \sigma_V, \mu), \sigma_1(\sigma_V, \mu)\} \Phi\{\ln V_{t+1}, \mu_1(V_t, \sigma_V, \mu), \sigma_1(\sigma_V, \mu)\} \times$$
$$\Phi\{\ln S_t - g(V_t|\Theta); \mu_2(S_{t-1}, V_{t-1}, \sigma_V, \theta_Z), \sigma_2(\theta_Z, \sigma_Z)\} \times$$
$$\Phi\{\ln S_{t+1} - g(V_{t+1}, t+1, \Theta); \mu_2(S_t, V_t, \sigma_V, \theta_Z), \sigma_2(\theta_Z, \sigma_Z)\}$$

We use non-informative and independent prior for Θ.

Because direct sampling from the above posterior distribution is impossible, the Metropolis-Hastings (MH) algorithm must be adopted. For proposal distribution (jumping distribution) we use,

For μ, $p_\mu(\mu^j|\mu^{j-1}) \sim Normal(\mu^{j-1}, 0.5)$.

For σ_V, σ_Z, $p_\sigma(\sigma^j|\sigma^{j-1}) \sim Gamma(c1, \sigma^{j-1}/c1)$.

For $\ln\theta_Z$, $p_{\theta_Z}(ln\theta_Z^j|ln\theta_Z^{j-1}) \sim Normal(ln\theta_Z^{j-1}, 1)$.

For V, $p_V(V_i^j|V_i^{j-1}) \sim Gamma(c2, V_i^{j-1}/c2)$.

where $c1$ and $c2$ are tuning parameters that control the step size.

However, updating one parameter at a time as described in this subsection, is not only computationally inefficient, but also becomes easily trapped in the wrong neighborhood of local extremum. To solve this problem, a simple Gibbs sampling method is adjusted by adopting the idea of simultaneous update.

1.4.2 Generalized Gibbs Sampling: Scale Transformation Update

To improve computational efficiency, we introduce an update that changes $\ln V_t$ and σ_V simultaneously because of the high correlation between $V_{1:T}$ and σ_V.

Given the current configuration of $(\mu, \sigma_V, \sigma_Z, \theta_Z, \ln V)$, the scale transformation proposes a move

$$(\ln V, \sigma_V) \rightarrow (\nu_1 \ln V, \nu_1 \sigma_V), \tag{1.25}$$

where ν_1 is scalar. In order to preserve the joint distribution, the ν_1 must be sampled from following distribution (Liu and Sabatti (2000), Kou, Xie, and Liu (2005)).

$$\begin{aligned} p(\nu_1) &\propto \nu_1^n P(\mu, \nu_1\sigma_V, \sigma_Z, \theta_Z, \nu_1 \ln V|S) \\ &\propto \nu_1^n f(\nu_1 \ln V, S|\mu, \nu_1\sigma_V, \sigma_Z, \theta_Z)\eta(\mu, \nu_1\sigma_V, \sigma_Z, \theta_Z). \end{aligned} \tag{1.26}$$

We implement the MH algorithm to sample from the density function described in equation (1.26). To propose ν_1, we use the gamma density $p_{\nu_1} = \Gamma(\nu_1; 1/c, c)$, which has a mean of 1. Then accept proposed ν_1 with the MH-like probability

$$\nu_1 = min\left(1, \frac{\Gamma(\nu_1^{-1}; 1/c, c)p_{\nu 1}(\nu_1)\nu_1}{\Gamma(\nu_1; 1/c, c)p_{\nu_1}(1)}\right). \tag{1.27}$$

Using the same reasoning, θ_Z and σ_Z are simultaneously updated as:

$$(\theta_Z, \sigma_Z) \rightarrow (\gamma_2\theta_Z, \gamma_2\sigma_Z). \tag{1.28}$$

The target distribution of γ_2 is

$$p(\gamma_2) \propto \gamma_2 P(\mu, \sigma_V, \gamma_2\sigma_Z, \gamma_2\theta_Z, \ln V|S). \tag{1.29}$$

1.4.3 Simulation Studies

Simple Gibbs Sampling Method vs. Generalized Gibbs Sampling Method

In order to assess the accuracy and efficiency of our approach, a simulation study is conducted. We simulate asset process V and market noise process Z with parameters $\mu = 0.1, \sigma_V = 0.1, \sigma_Z = 0.1$, and $\theta_Z = 30$. The parameter values are chosen in a way that is consistent with the real data. We use the median values (rounded) obtained from real data (eight firms which will be described in section 1.5). Stock process S is generated based on the equation (1.2). 10000 samples are drawn from the posterior distributions described in equation (1.22) and equation (1.23).

To prevent an explosion of the estimated θ_Z, constraints to a noise process Z are applied as follows:

$|Z_t| < 0.1$

$\theta_Z < 100$ and positive.

First, we assume there is at most 10% difference between market and theoretical (derived from the Black-Cox pricing model) equity values. This is a fairly generous assumption compared to the previous results in Duan and Fulop (2009).[3] Second, the mean-reverting rate of noise is assumed to be positive and less than 100. Only with the short term data (6 months in this essay), separate identification of mean-reverting rate and volatility of the noise process is difficult, so a constraint is imposed on mean-reverting rate. In the presence of a large mean-reverting rate, there is a marginal gain of adopting a mean-reverting process instead of independent normal distribution; thus θ_Z is assumed to be less than 100.

[3] According to Duan and Fulop (2009), the mean of standard deviations of stock market noise of 100 randomly chosen non-DowJones30 companies is 0.0043 and the 90 percentile is 0.016 under Merton's structural model assumption. The mean of standard deviations of stock market noise of DowJones30 companies is 0.003 and the 90 percentile is 0.007.

The simulation results are then summarized with histograms, autocorrelations, and plots of posterior samples. The fast decay of the autocorrelations suggests a speedy convergence of the algorithm. Figures 1.2 and 1.3 are estimation results without using the scale update (the simple Gibbs sampling method). Figures 1.4 and 1.5 are estimation results with the scale update (the generalized Gibbs sampling method).

The simulation results tell us that by adopting the scale update (the generalize Gibbs sampling), the algorithm converges faster for all parameters. Moreover, we can identify all the parameters more correctly than by using simple Gibbs sampling method. The estimation performance of σ_Z is significantly improved. For the results of the simple Gibbs sampling, the sampled value is trapped in an over-estimated value, while for the results of the generalized Gibbs sampling, the sampled σ_Z is near the true value 0.1. However, problems still remain in estimating θ_Z. We might need to consider different re-parametrizations or further scale transformations to solve this problem. Figure 1.6 provides details for the estimating the hidden asset process using the scale update.

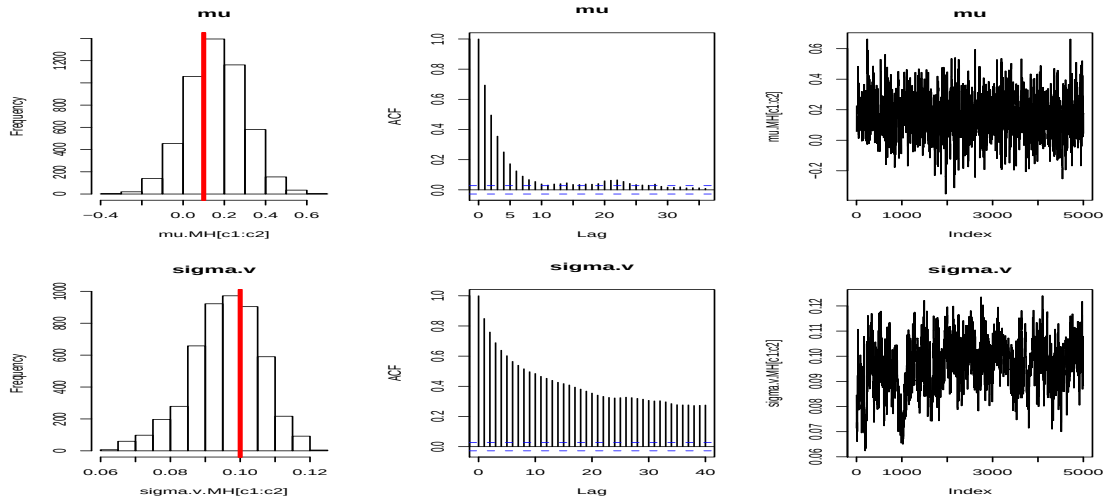

Figure 1.2: **Asset Process Parameters Estimation Results without the Scale Update (in Simulated Data Set).** The top panel is μ and the bottom panel is σ_V. From left, histograms of the posterior samples (| : true value), autocorrelation of posterior samples and plot of posterior samples.

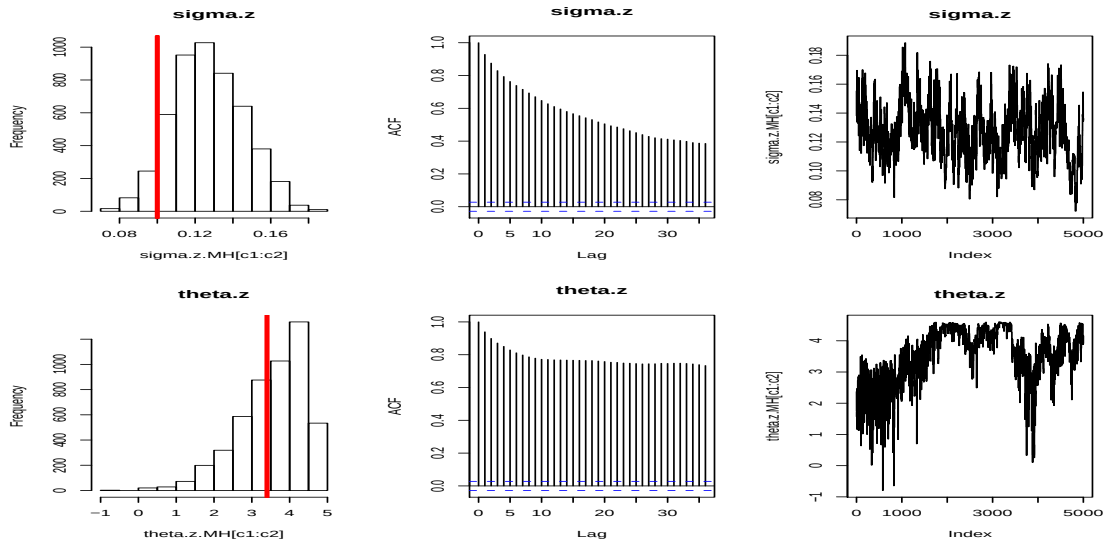

Figure 1.3: **Noise Process Parameters Estimation Results without the Scale Update (in Simulated Data Set).** The top panel is σ_Z and the bottom panel is $\ln\theta_Z$. From left, histograms of the posterior samples (| : true value), autocorrelation of posterior samples and plot of posterior samples.

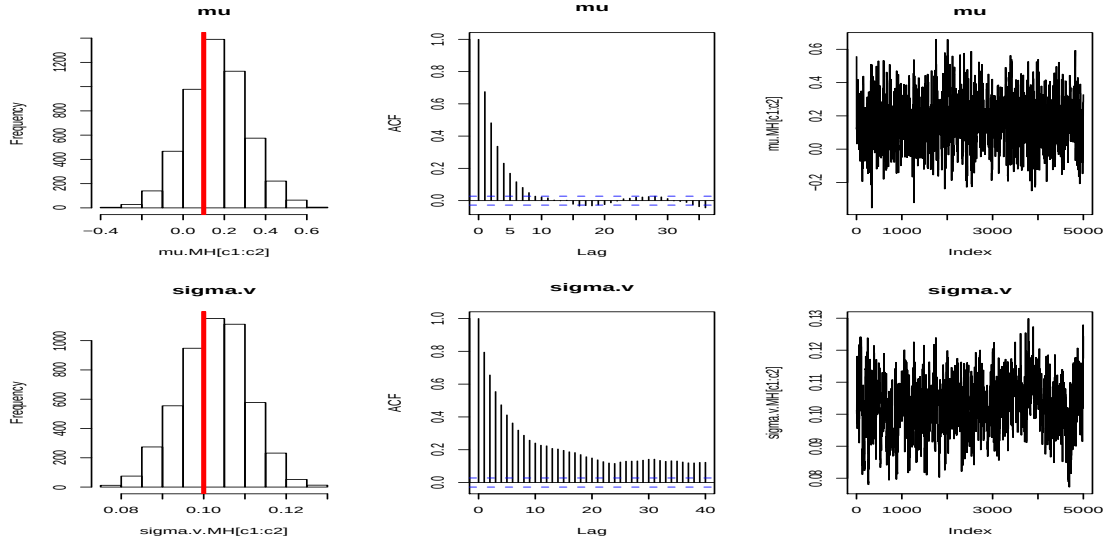

Figure 1.4: **Asset Process Parameters Estimation Results with the Scale Update (in Simulated Data Set).** The top panel is μ and the bottom panel is σ_V. From left, histograms of the posterior samples (| : true value), autocorrelation of posterior samples and plot of posterior samples.

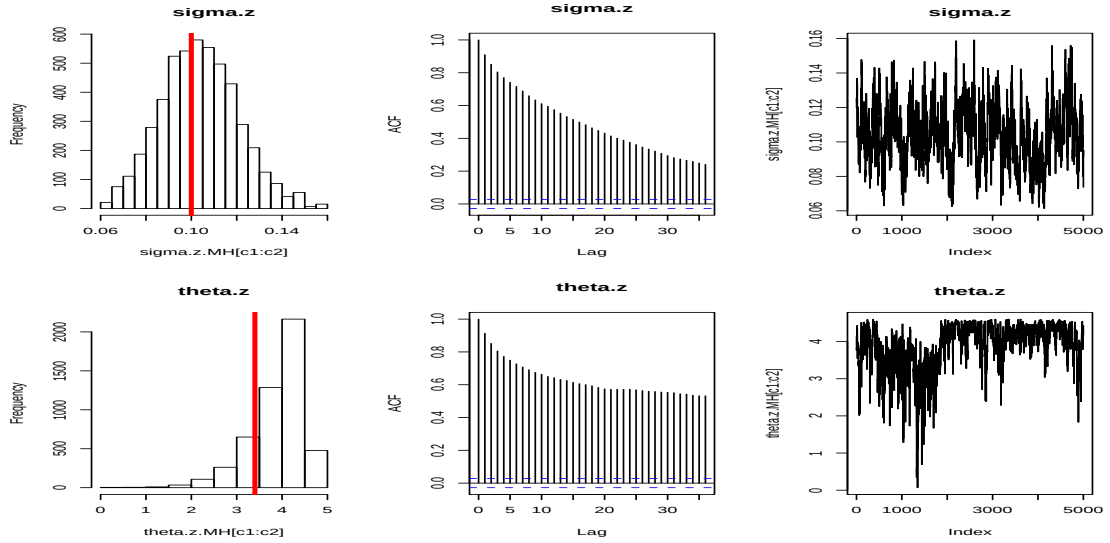

Figure 1.5: **Noise Process Parameter Estimation Results with the Scale Update (in Simulated Data Set).** The top panel is σ_Z and the bottom panel is $\ln\theta_Z$. From left, histograms of the posterior samples (| : true value), autocorrelation of posterior samples and plot of posterior samples.

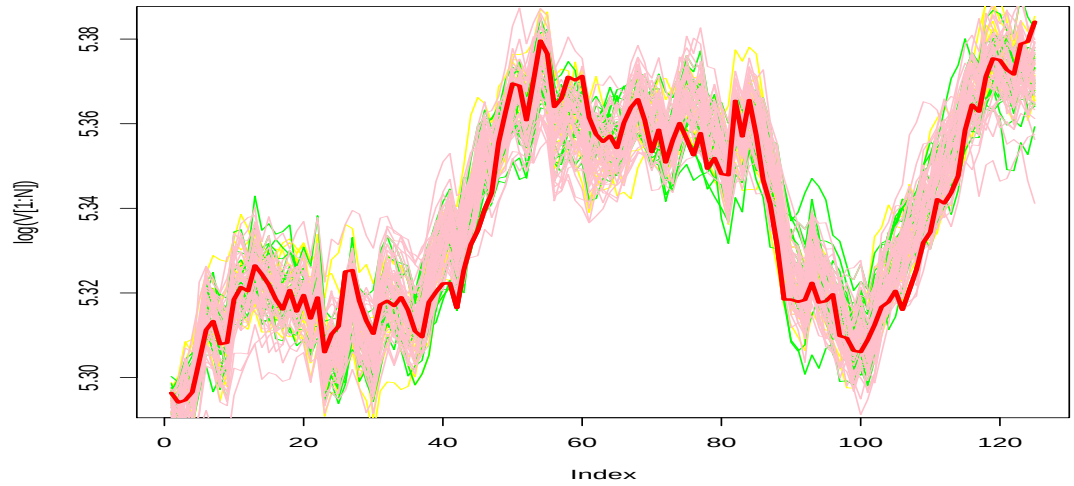

Figure 1.6: **Asset Process Estimation with Scale Update (in Simulated Data Set).** True $\ln V$ process is marked with red bold line. Yellow lines are posterior sample paths between iteration number 1000 to 3000, green lines are between 3000 to 5000, and pink lines are between 5000 to 10000.

Small Volatility vs Large Volatility

We found that the generalized Gibbs sampling method offers improvements in the parameter estimation performance. To investigate more thoroughly the parameter estimation performance, another data set is simulated with different parameter configurations. We vary the three volatility parameters, σ_V, σ_Z, and θ_Z, to investigate their effect on performance. The parameter values used are $\mu = 0.1, \sigma_V = 0.2, \sigma_Z = 0.2, \ln\theta_Z = 50$, which are close to the 90 percentile of the estimates obtained from real data (which will be described in section 1.5).

Tables 1.3 and 1.4 show the summary statistics of posterior samples using the generalized Gibbs sampling method for the two different simulated data sets. Table 1.3 reports the results when the volatility parameters are small: $\mu = 0.1, \sigma_V = 0.1, \sigma_Z = 0.1, \theta_Z = 30$. Table 1.4 reports the results when volatility in asset and market noise processes are all larger than previous data set: $\mu = 0.1, \sigma_V = 0.2, \sigma_Z = 0.3, \theta_Z = 50$. In terms of the error between posterior mean and the true values of volatility parameters, σ_V and σ_Z are identified more correctly in the case with smaller values.

Table 1.3: **Small Volatility Estimation Results.** This table summarized the estimation Results when data set are simulated with the true parameters: $\mu = 0.1, \sigma_V = 0.1, \sigma_Z = 0.1, \theta_Z = 30$.

Parameter	Posterior Mean	True value	Posterior s.d.	Error	Posterior s.d/ True
μ	0.166	0.1	0.146	0.066	1.46
σ_V	0.102	0.1	0.008	0.002	0.08
σ_Z	0.104	0.1	0.017	0.004	0.17
$\ln\theta_Z$	3.890	3.401	0.598	0.490	0.176

Table 1.4: **Large Volatility Estimation Results.** This table summarized the estimation results when data set are simulated with the true parameters: $\mu = 0.1, \sigma_V = 0.2, \sigma_Z = 0.3, \theta_Z = 50$.

Parameter	Posterior mean	True value	Posterior s.d.	Error	Posterior s.d/ True
μ	0.068	0.1	0.297	-0.032	2.97
σ_v	0.211	0.2	0.012	0.011	0.06
σ_z	0.277	0.3	0.054	-0.023	0.18
$\ln\theta_Z$	3.72	3.91	0.63	-0.190	0.16

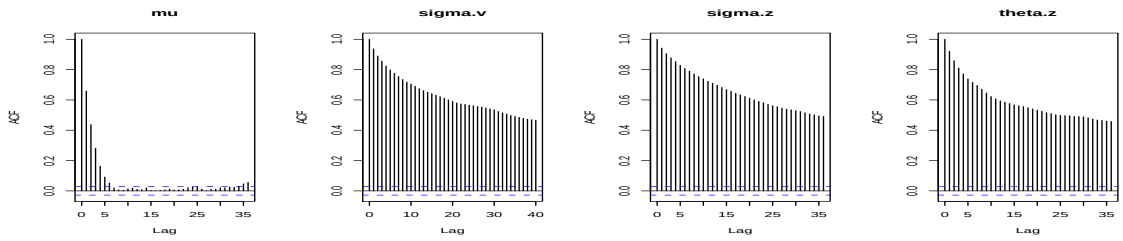

Figure 1.7: **Convergence of Posterior Sample in Large Volatility Case.** These graphs are ACFs of posterior samples of (from the left) μ, σ_V, σ_Z and θ_Z when the true parameters are $\mu = 0.1, \sigma_V = 0.2, \sigma_Z = 0.3, \theta_Z = 50$.

Figure 1.7 shows autocorrelation of posterior samples for $\mu, \sigma_V, \sigma_Z, \theta_Z$ in the case of large volatility. Compared to the Figures 1.4 and 1.5, convergence of the algorithm for larger volatilities is not as good as for smaller volatilities.

1.5 Empirical Analysis

Using the method developed in Section 1.4, we analyze empirical market data. The generalized Gibbs sampling method is implemented for parameter estimation of eight different companies. The first four companies, 3M, AT&T, IBM, and Coca Cola Company constitute the Dow Jones Industrial Index. The other four companies, Comcast, Sunoco, Washington Post, and Time Warner, Inc. do not belong to the Dow Jones Industrial Index. The reason for selecting companies in these two different groups is that the Dow Jones companies tend to have more heavily traded stocks than others. By choosing firms from both DowJones and non-DowJones companies, we can diversify our sample with small size.

The data set consists of daily equity prices for 6 months, from January 3, 2007 to July 3, 2007. The closing prices of equity and the numbers of outstanding are taken from the CRSP database. The balance sheet information is from the Compustat annual file. The product of the closing prices and the numbers of outstanding is considered as observed equity values.

For the settings of values, which are assumed to be exogenously given in the model, we follow Duan and Fulop (2009). For risk-free interest rate r, the one-year Treasury constant maturity rate

Table 1.5: **Estimation Result for Empirical Data, By Firms.** This table show the posterior means and posterior standard deviation (in parenthesis) of parameter μ, σ_V, σ_Z, and $\ln \theta_Z$.

Company	$\hat{\mu}$	$\hat{\sigma}_V$	$\hat{\sigma}_Z$	$\hat{\ln\theta_Z}$
3M	0.2(0.12)	0.08(0.003)	0.12(0.17)	2.7(0.93)
AT&T	0.72(0.48)	0.33(0.018)	0.28(0.06)	3.8(0.63)
IBM	0.17(0.18)	0.12(0.006)	0.10(0.018)	1.0(1.29)
Coca Cola Company	0.14(0.13)	0.08(0.004)	0.08(0.013)	3.45(0.75)
Comcast	-0.04(0.23)	0.16(0.008)	0.17(0.03)	3.9(0.58)
Sunoco	0.2(0.26)	0.19(0.009)	0.15(0.03)	3.5(0.89)
Washington Post	0.05(0.13)	0.09(0.005)	0.07(0.013)	3.3(1.09)
Time Warner INC.	-0.04(0.194)	0.135(0.006)	0.09(0.017)	3.86(0.59)

obtained from the U.S. Federal Reserve, is used. The initial maturity of debt is set to two years. For the face value of the bond, the book value of liabilities of a company at the end of 2006 is compounded for two years at the risk-free interest rate r. Other initial values are set as followings: V_0=S(at Dec,29,2006)*(1-market leverage), γ=0.1, and $L_0 = 0.8 * V_0$.

Table 1.5 reports estimation results for all eight firms. In general, DowJones companies have a higher mean level in asset process. The mean-reverting rate of noise process is higher in non-DowJones companies than in DowJones companies, in general. In discrete time settings, the AR(1) coefficient is $\exp(-\theta_Z \Delta t)$. In DowJones companies, it ranges from 0.837 to 0.989 and in non-DowJones companies, it ranges from 0.822 to 0.898. The mean of market noise at time t is closer to its previous values in Dow Jones companies.

To ascertain whether our noise estimates are in line with empirical findings, we conduct a cross-sectional analysis on market noise in relation to the commonly adopted proxies for market liquidity: trading volume and bid-ask spread (Fleming (2003)). A negative relationship between market noise and trading volume and a positive relationship between market noise and the bid-ask spread are expected (Ait-Sahalia and Yu (2009)). Trading volume and percentage bid-ask spread at January 3, 2007 are used as proxies for market liquidity during data period. The trading volume is the daily volumes from the CRSP daily file. For the percentage bid-ask which is bid-ask spread divided by bid price, we use CRSP daily files to get the closing ask and bid. As a measure of the size of market noise, we use the relative ratio between posterior mean of asset $\hat{\sigma}_V$, and posterior

Table 1.6: **Estimated Market Noise and Proxies for Market Liquidity By Firm.** As a measure of the size of market noise, we use the relative ratio between posterior mean of asset $\hat{\sigma}_V$, and posterior mean of trading noise volatility $\hat{\sigma}_Z$. Trading volume and percentage bid-ask spread at January 3, 2007 are used as proxies for market liquidity during data period. The trading volume is the daily volumes from the CRSP daily file. For the percentage bid-ask which is bid-ask spread divided by bid price, we use CRSP daily files to get the closing ask and bid.

Company	$\hat{\sigma}_Z/\hat{\sigma}_V$	Trading Volume	Percentage Bid-Ask Spread
3M	1.5	3,781,700	0.090
AT&T	0.848	33,710,000	0.029
IBM	0.833	9,199,500	0.041
CoCa Cola Company	1	7,877,300	0.103
Comcast	1.06	13,262,374	0
Sunoco	0.789	4,702,900	0.050
Washington Post	0.777	24,500	0.102
Time Warner Inc.	0.666	27,394,400	0.045

mean of trading noise volatility $\hat{\sigma}_Z$. A firm with large $\hat{\sigma}_Z/\hat{\sigma}_V$ is expected to have more volatile market noise. Table 1.6 shows those values in each firm.

.

Correlation between trading volume and $\hat{\sigma}_Z/\hat{\sigma}_V$ is -0.346. Correlation between percentage bid-ask spread and $\hat{\sigma}_Z/\hat{\sigma}_V$ is 0.234. These two correlation results indicate that the magnitude of our estimated values are plausible.

1.6 Conclusion

In this essay, we built the structural credit risk model and proposed an estimation method under certain assumptions, which are relaxed and closer to the real market. Under the Black-Cox model, asset process and model parameters estimation are implemented when the stock market is contaminated with mean-reverting processed market noise. Moreover, the cross-asset class research is performed on sequential hidden asset process estimation, given all parameters. For estimation, two different Bayesian methods are adopted: the particle filter algorithm and the generalized Gibbs sampling method.

First, we sequentially estimated hidden asset values using the particle filter algorithm. In the sense of cross-asset class research, we showed how much estimation performance is improved by

adding option market information. The simulation results showed that estimation using both stock and option price information produced a smaller MSE with a smaller posterior standard deviation and a shorter prediction interval on average than estimation using only stock price information. These results imply that improvement in accuracy and efficiency can be achieved in estimation by adding related cross-asset class information.

Second, we adopted a generalized Gibbs sampling method to estimate asset process and model parameters together. Our simulation study shows that the generalized Gibbs sampling method estimates parameters successfully, especially on asset model parameters μ and σ_V. In addition, by adopting the generalized Gibbs sampling algorithm, we were able to enhance the convergence rates of all model parameters estimates. We also applied the methodology to real data. Using the signs of correlation coefficients between the relative size of market noise and two proxies for market liquidity, we showed how our estimates accurately reflect real market properties in terms of volatilities. The method used in this essay can be applied to more complicated models, particularly when closed-form solutions are difficult to obtain.

The model we proposed in this essay can be easily extendable in terms of incorporating more varied securities information. Therefore, as future works, we plan to incorporate other types of derivatives. In sequential hidden asset value estimation, additional information (option price information) enhanced performance, but could not provide substantially large improvements in estimation. This is a reason for not incorporating option price information in simultaneous estimation of parameters and asset process. We expect that including credit default swap rates will achieve substantial improvements in the model. Credit default swap rates are often better indicators of default risks than corporate bonds since corporate bond prices may reflect tax effects and liquidity risks, other than default risks.

2. INDUSTRY-SPECIFIC CORRELATED DEFAULTS

2.1 Introduction

Historically, there are periods of time when only certain industries suffer greatly. Despite a favorable overall economy, twenty-two companies in the oil or oil services industry defaulted on rated debt between 1982 and 1986 (Lucas (1995)). More recently, during the recent financial crisis, large financial institutions have collapsed. AIG, Bear Stearns, Fannie Mae and Freddie Mac needed government bailouts or takeovers to survive. Lehman Brothers is in bankruptcy. Merrill Lynch has been sold. Does a crisis in a certain industry affect equally or at least similarly all industries? If it does not, how much do the effects differ? Our study on industry-specific correlated default starts from these questions. In this essay, the model for the default probability is extended from previous default correlation studies in order to include industry-specific correlation effects.

There is considerable evidence to show that defaults have occurred in distinct clusters across a given time frame (see Figure 2.1). In order to explain this phenomenon, studies on the correlated defaults have started. Duffie and Garleanu (2001) first developed the idea of common factor to explain the default clusters. They used the intensity credit risk model based on the assumption that firms may be exposed to common or correlated risk factors whose co-movements cause change in conditional default probabilities that are correlated across firms.

As in Duffie and Garleanu (2001), under the intensity model setups, default correlation studies have been developed by relaxing the conditional independence assumption and allowing default intensities to depend on common factors. Duffie and Wang (2007) provided maximum likelihood estimators of term structures of conditional probabilities of corporate default by incorporating the dynamics of firm-specific and macroeconomic covariates. In contrast, under the structural

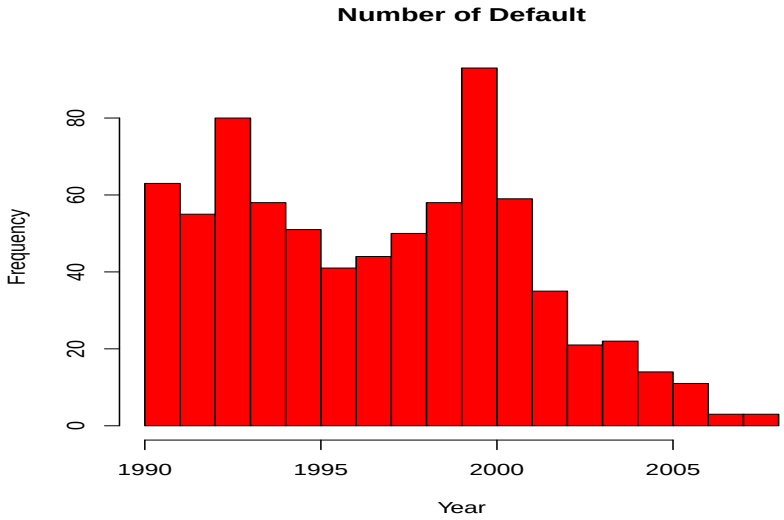

Figure 2.1: **The Number of Defaults between 1990 and 2008.** This histogram is based on U.S. public firms and shows the number of defaults in each year.

model setups, default correlation research adopted the direct modeling of the joint dynamic at an underlying asset level. Zhou (2001) expanded the standard structural model of Merton (Merton (1974)) to two correlated asset processes.

In modeling multi-firm defaults and their default dependence, the typical intensity model has several advantages over the typical structural model in terms of correlation measuring method. The typical intensity model has applied the factor-structure method as in Duffie and Garleanu (2001) while typical structural model has directly measured the asset correlation as in Zhou (2001). One of the advantages of a factor-structured method over direct measurement methods is that a large number of firms can be modeled at the same time and the model can be easily extendable to additional firms. Moreover, parameters related to the correlation can be estimated easily and more effectively using the factor-structured method.

In this essay, we build a default correlation model with a factor structure on default intensities as in Duffie and Wang (2007). However, we extend previous studies by adopting industry-specific hidden factors. The concept of a hidden common factor was first introduced in Duffie, Eckner, Horel, and Saita (2009). They found that, after controlling for observed covariates, defaults were persistently higher than expected during a lengthy period of time. They suspected the existence

of missing covariates which have an important role in explaining the default correlation. In order to adjust this bias in their previous analysis (Duffie and Wang (2007)), they added the hidden common factor called *frailty*.

However, with this method, bias in default distribution estimation can be partially adjusted because time-dependent correlation path is assumed to be same for all firms. Not all dependence patterns are the same for all firms. For example, the default correlation between Exxon and Chevron could be very different from that between Exxon and Wal-Mart. The default correlation between firms in the same industry can be higher than for firms in different industries. In fact, Moody's Binomial Expansion Technique (BET) is based on the idea that firms in the same industry are related, while firms in different industries could be treated as independent. Furthermore, within-industry correlations and between-industries correlation are not the same for all industries.

With Credit default swap (CDS) spread, we can give more concrete rationale on the not-all-same within- and between- industries default correlation. It is easiest to think of CDS spread as an insurance fee for the contract ensuring protection against default, thus the value of CDS spread of a firm is known as a good indicator of its default risk. In fact, much research on defaults test model performance with CDS prices. Table 2.1 is the correlation matrix of 5-year CDS prices during 2009, in each subcategory of industry. Banks and Insurance, which are both subcategories of the financials industry, correlated more than Banks and Restaurant; Restaurant is a subcategory of the consumer discretionary industry. The within-industry CDS correlations are also different in each industry. For example, the correlation between Restaurant and Leisure, which are both subcategories of consumer discretionary, is much smaller than Banks and Insurance. The between-industry correlations are also different. The correlation between Bank and Telecom is much larger than Bank and Leisure or Bank and Restaurant.

Real market data (CDS prices) suggest that the default risk correlations are not all the same in 2009. Even though, it is possible to change the relative size and/or order of correlation coefficients, the fact that correlation coefficients are different industry-specifically might be continued in other time periods. De Servigny and Renault (2002) also showed inequality between within- and between-industry correlations by providing default correlations for non-investment-grade bonds

Table 2.1: **Correlation between CDS Prices in Each Sector.** These correlation coefficients are derived from daily 5-year sector-specific CDS prices between May 1, 2009 and December 31, 2009. Dates with missing flags are deleted, so we use 160 days of data. Data source is Bloomberg.

	Banks	Insurance	Leisure	Restaurants	Retail	Telecom
Banks	1.00	0.95	0.91	0.19	0.89	0.67
Insurance	0.95	1.00	0.97	0.24	0.90	0.64
Leisure	0.91	0.97	1.00	0.31	0.92	0.68
Restaurants	0.19	0.24	0.31	1.00	0.29	0.72
Retail	0.89	0.90	0.92	0.29	1.00	0.61
Telecom	0.67	0.64	0.68	0.72	0.61	1.00

by sectors. They used simple default probability and joint probability estimates as given by Lucas (1995). However, we propose a model based on the factor-structured method which makes it possible to measure default correlation more effectively than De Servigny and Renault (2002).

In this essay, we incorporate these different industry correlations into our model by extending the idea of *frailty* in Duffie, Eckner, Horel, and Saita (2009). By assuming industry-specific factors as hidden and estimating them through the data set, different industry factors can be fairly incorporated. Because it is hard to find corresponding industry-specific variables that have long history and represent default risk, the hidden factor approach suggests a way to incorporate an industry factor into the model.

More than incorporating industry-specific factors into the model, the contributions of this essay are as follows. By assuming an industry hidden factor as a random effect, within- and between-industry correlations are able to be measured and hence to be compared. We also assign a time-series model on these random effects and estimate all parameters, thus making prediction possible. Based on monthly data encompassing 19 years for the 13,682 U.S. public firms (which is 1,413,314 firm-months of data), we compare the relative scale of within- and between- industry correlations. Through empirical results, we also show that industry-specific factors resolve the over- or under-estimation of default probabilities.

The basic methodology used in this essay is Bayesian statistical estimation. The Monte Carlo Expectation and Maximization (MCEM) method is used to estimate the random effects (hidden factors) and parameters in the model. MCEM is a good choice as a statistical estimation method

because it takes into consideration the fact that there are missing (hidden) covariate paths and that the E-step is analytically difficult to compute.

This essay is organized as follows. Section 2.2 gives an overview of data and defaults in each industry. Section 2.3 introduces our default intensity model with observed and hidden common factors and in Section 2.4, details are provided on how unobserved hidden factors are modeled to capture within- and between- industry correlations. In Section 2.5, the estimation procedure is described. Section 2.6 shows the empirical results and comparison of estimation performance to previous models in section 2.7. Finally, the essay ends with discussion for further studies in Section 2.8.

2.2 Data: Default and Industry

This section introduces the data and provides a summary of them aimed at motivating our approach. With a simple summary of the data set, we show the necessity of industry-specific analysis in default correlation study.

2.2.1 Data Overview

The data set in our study contains $1,413,314$ firm-months of data from January 1990 to December 2008. Data for $13,682$ U.S. public firms taken from Compustat and CRSP are used. For identifying firms between two data sources, a CRSP/Compustat Merged Database (CCM) from the Wharton Research Database Service (WRDS) is used. Because public-firm defaults levels are relatively low, we rely on 19 years of data from 1990 to 2008 to make sure that suitably many defaults are included to conduct statistical analysis.

Only those firms that have at least 6 months of information beginning January 1990 are used. We consider the research company deletion date (data item DLDTE) as the last date of the firm and ignore all other variable values after that date. Firms without industry categories (data item GSECTOR) or without exit reason (data item DLRSN) are excluded from the final data set.

All balance sheet information is taken from quarterly Compustat files. For cases with missing debt values, we replace these with the nearest value in the same year. For cases where an entire

Table 2.2: **Number of Firm Exits of Each Type from 1990 to 2008.**

TYPE	Number
Active (Not Exit)	5,119
Bankruptcy	503
Other Default	258
Merge & Acquisition	5,913
Other Exit	1,889

year's values are missing, we replace the values with values from yearly Compustat files. If yearly values are also missing, we then treat these as final missing values.

Table 2.2 shows the number of firms in each exit category. Of the 761 total defaults, 503 first occurred as bankruptcies, although many of the "other defaults" eventually led to bankruptcies. For the construction of a data set and for obtaining the exact definition of these event types, we follow the description given by Duffie and Wang (2007).

2.2.2 Industry Categories

Companies are categorized based on the Global Industry Classification Standard (GICS). The GICS is the industry classification structure used for Standard & Poor's U.S. industry index calculations (Maitland and Blitzer (2002)). GICS is currently comprised of 10 sectors and 23 industry groups. The firms are grouped based on 10 sectors and combines four sectors of materials, consumer staples, telecommunications services and utilities all of which contain a relatively small number of companies. We will hence forth use the term *industry* instead of *sector* to refer to the following seven categories:

1. Consumer discretionary - automobiles & components, consumer durables & apparel, hotels, restaurants & leisure, media, retailing.

2. Energy - energy equipment & services, oil & gas.

3. Financials - banks, diversified financials, insurance, real estate.

4. Health care - health care equipment & services, pharmaceuticals & biotechnology.

5. Industrials - capital goods, commercial services & supplies, transportation.

6. Information technology - software & services, technology hardware & equipment.

7. Others (materials, consumer staples, telecommunications services, utilities sectors).

2.2.3 Defaults in Each Industry

Table 2.3 shows the number of defaults in each industry from 1990 to 2008. Generally, 3 to 5% of firms in most industries default. The industry with the largest default rate is the consumer discretionary industry, with a 10% default rate. Except for this number, there are no major differences across industries in the default rate, spanning 19 years.

Table 2.3: **Number of Defaults in Each Industry from 1990 to 2008.**

Industry	Consumer	Energy	Financials	Health	Industrials	Infotech	Others
# of Firms	2423	633	3078	1696	1771	2564	1517
# of Defaults	247	20	121	64	108	132	69

In contrast, the default trends of each industry are different from each other, as shown in Figure 2.2. The most notable industry with a unique pattern is the energy industry. While in other industries, default rates are comparatively high near 2000 as well as 1992, in the energy industry, default rates in the early 1990s is very high but in 2000 there is only very little increase. In the financials industry, defaults occurred chiefly in 1995, when in most other industries, there were only a few firms that defaulted immediately after high-default years.

It seems there are fairly large differences across the industry but we can still find some similarities among them. In general, around 1992 and around 2000 were high peak default years in all industries and after 2005 default rates are stabilized in low values in all industries. Moreover, among the different patterns, we can find the pair or group of industries that show relatively more consistent patterns. For example, the default rate patterns observed before and after 2000 can be classified into two groups. In the consumer discretionary and industrials industries, most defaults happened before 2000 and later the default rate decreased rapidly and then stabilized. However,

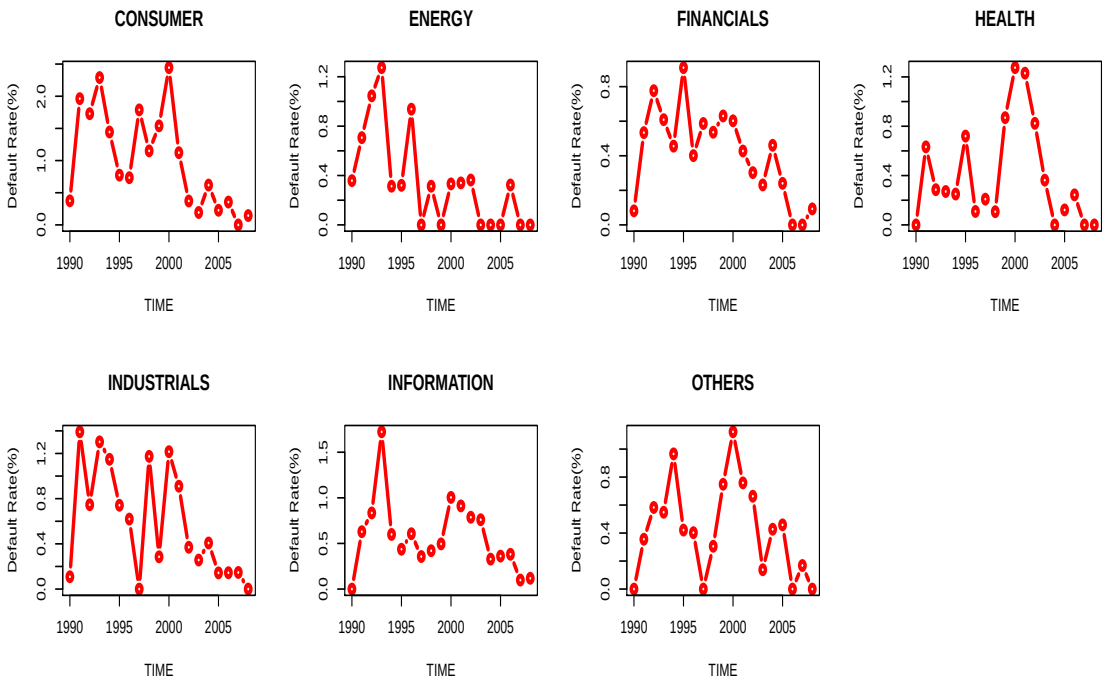

Figure 2.2: **Yearly Default Rate in Each Industry.** These graphs show (the number of defaults)/(the number of firms) in each industry for each year between 1990 and 2008.

for firms in the financials and information technology industries, the default rate decreased slowly after 2000.

As Duffie and Garleanu (2001) assumed the existence of a common factor path determining the time cluster of defaults events, we also explain industry-specific time cluster of defaults with an industry-specific factor in this essay. Instead of one common factor for all the firms, we assign the same factor process for the firms only within same industry, thus their default probabilities are correlated by sharing the same industry-specific factor process. However, as mentioned before, industry-specific common factors are not all completely different from each other. The existence of common features among them means they are connected with each other, but the degree of connectivities are different pair-wisely or group-wisely. For example, defaults in the consumer discretionary industry might be more correlated with defaults in the industrials industry than with other industries (such as financials and information technology). Similarly, defaults in the financials industry might be more correlated with defaults in the information technology industry than with any other.

The default data summaries thus provide us sufficient information to establish a model for industry-specific default analysis. As shown earlier, not only is there a correlation of defaults within an industry, but defaults across industries can also have a correlation. Furthermore, the level of such a correlation is different for each pair or groups of pairs. In this essay, we thus propose a correlated industry-specific hidden factor to capture the industry-specific, time-varying correlation structure. In the following section, we discuss our model in more details.

2.3 Model

In the previous section, the necessity for industry-specific analysis was established by exploring default cluster patterns. Here we expand on the frailty concept described in Duffie, Eckner, Horel, and Saita (2009) to industry-specific terms. This section provides a detailed specification for a model on default intensity. The model is first introduced, followed by a discussion on the observed and hidden common factors adopted in this essay.

2.3.1 Mixed Effect Model with Correlated Industry-Specific Default Intensity

As for the typical intensity model, we suppose default intensities of all firms at time t depend on a state vector X_t and that, given the path of the state vector process X, the default times for different firms are conditionally independent. By assuming that all firms' default intensity share the same X_t at each time t, we can incorporate a correlation structure for entire companies into the model. We assume that, conditional on the path of X determining the default intensities, the default times of firms are the first event times of an independent Poisson process with time-varying intensities determined by the path of X.

The hidden factor idea proposed in Duffie, Eckner, Horel, and Saita (2009), however, departs in an important way from the traditional setting by assuming that X is not fully observable to the econometrician. Similarly, we adopt an industry-specific hidden factor into our model. In the financial market it is difficult to find the coordinate variables related to each industry. For this reason, we define an industry-specific factor as a hidden factor. By doing so, not only can we evenly incorporate industry factors into our model, but we can also capture unobservable risk

factors into our model.

In order to carry out the statistical analysis, we discretized time as $\Delta t = 1/12$ (monthly data). The probability that the i^{th} firm's default event happens during the time interval $[t-\Delta t, t]$ is

$$P(D_{i[j],t}) = \lambda_{i[j],t}\Delta t. \tag{2.1}$$

Under the factor structure, we model the conditional mean arrival rate of default, $\lambda_{i[j],t}$, with a mixed effect model, as given below:

$$\lambda_{i[j],t} = \exp\{a_0 + a_1' U_{i,t} + a_2' X_t + Y_{j,t}\}, \tag{2.2}$$

$$\vec{Y_t} = K\vec{Y_{t-\Delta t}} + \vec{\varepsilon_t}, \tag{2.3}$$

$$\varepsilon_t \sim N(0,\Sigma)_J, \tag{2.4}$$

where

$$\vec{Y_t} = (Y_{1,t}, Y_{2,t}, ..., Y_{J,t})'$$

$$\vec{\varepsilon_t} = (\varepsilon_{1,t}, \varepsilon_{2,t}, ..., \varepsilon_{J,t})'.$$

The index i refers to the firm i, where $i = 1, ...I$ and the index j is for industry, where $j = 1, ...J$. t is for time, where $t = 1, ...T_i$ and T_i is min(Exit time of i^{th} firm, T), where T is the data end-period. $U_{i,t}$ denotes the vector of variables that are specific to i^{th} firm, and X_t denotes common macroeconomic variables vector. $Y_{j,t}$ denotes j^{th} industry-specific factor at time t. We use the index $i[j]$. This means that the i^{th} firm belongs to the j^{th} industry, where there are J industries in total. We use the firm's distance to default and trailing one-year stock return as firm-specific variables U_{it} and the three-month Treasury bill rate and the trailing one-year return on S&P500 index as common macroeconomic variables X_t. a_1' and a_2' are corresponding coefficient vectors. More details on these covariates will be described separately in following two subsections.

Our model is summarized as a mixed effect model with the following random and fixed effects.

1. Random effect (Hidden factor effect)
    a. Y_{jt}: j^{th} industry effect at time t.
2. Fixed effect (Observed factor effect)
    a. a_1: Effect of firm-specific variables.
    b. a_2: Effect of macroeconomic variables.

The description of each effect is provided in the next two subsections.

2.3.2 Observed Factors: $U_{i,t}, X_t$

The observed factors used in this essay are again divided into two categories. One is a firm-specific factor ($U_{i,t}$), and the other is a macroeconomic factor (X_t). We follow Duffie, Eckner, Horel, and Saita (2009) for selecting these factors as shown below:

1. Firm-specific variables ($U_{i,t}$)
    1.1. Distance to default (DTD) : Distance to default means the distance of a firm's asset growth from its liability. This variable is not directly observed, so it needs to be calculated along with other financial variables as follows:

$$DTD_t = \frac{ln(V_t/L_t) + (\mu_A - 0.5\sigma_A^2)T}{\sigma_A\sqrt{T}}, \tag{2.5}$$

    where V_t is the firm's market asset value, L_t is its liability and μ_A, σ_A are the firm's mean rate of asset growth and asset volatility. Because asset value is not an observable variable in the market, we construct this with an iterative method described in Duffie and Wang (2007). In order to construct distance to default, we use stock prices, number of outstanding, short-term debt, long-term debt, and the one-year Treasury-bill rate.
    1.2. The firm's trailing one-year stock return (Lag12) : This covariate suggested by Shumway (2001).

2. Macroeconomic variables (X_t)

    2.1. The three-month Treasury bill rate (Tbill3): This covariate plays a role as an effect of monetary policy.

    2.2. The trailing one-year return on the S&P 500 index (SP500): This covariate measures the market return.

Duffie and Wang (2007) and Duffie, Eckner, Horel, and Saita (2009) give a detailed description of these covariates and discuss their relative importance.

2.3.3 Hidden Factor ($Y_{j,t}$)

The need and importance of the hidden factor in a default model are discussed in several recent studies. Vasicek (1991) and Gordy (2003) showed that defaults are more heavily clustered in time than currently captured in models with observed covariates. Duffie, Eckner, Horel, and Saita (2009) found that, after controlling for observed covariates, defaults were persistently higher than expected during a lengthy period of time from 1986 to 1991 and persistently low during the mid-90s. To capture this un-captured, time-varying default correlation, Duffie, Eckner, Horel, and Saita (2009) introduced the idea of a common, dynamic, latent factor named *frailty*, which was driving the default.

By extending the frailty model, we set the industry-specific hidden factor as a random effect with a probability distribution. Different time patterns of defaults (Figure 2.2) show that an industry-specific default analysis is therefore required. However, this is not enough evidence for the necessity of incorporating industry-specific hidden factors, because we also use the firm-specific variables for the intensity model. It is possible that these firm-specific observed variables are enough to resolve industry-specific default time patterns. We thus conduct a simple test with the residuals of the Generalized Linear Model (GLM).

First we fit following GLM only with observed factors as predictors:

$$\begin{aligned} P(D_{i[j],t}=1) &= \lambda_{it}\Delta t, \\ \lambda_{i[j],t} &= \exp\{a_0 + a_1 U_{i,t} + a_2 X_t\}, \end{aligned} \tag{2.6}$$

where $D_{i[j],t}$ is the default indicator of i^{th} firm which belongs to j^{th} industry, between the time interval $[t-\Delta t, t]$. We then extract the following Pearson residuals:

$$Presi_{i[j],t} = \frac{D_{i[j],t} - \hat{\lambda}_{i[j],t}}{\hat{\lambda}_{i[j],t}(1-\hat{\lambda}_{i[j],t})}. \tag{2.7}$$

Figure 2.3 shows the mean of Pearson residuals for each year within each industry. Even after controlling all observed factors, the energy industry shows a fairly unique pattern compared with others. The health care and the "others" industries show similar Pearson residual patterns. Other than these two, all residual patterns are considerably different from each other. These different residual series imply that the different default time patterns are not resolved even after controlling firm-specific factors; therefore additional industry-specific information is needed.

This section provides a discussion on our industry-specific hidden factor model and common factors that will be applied to our model. A simple preliminary study based on Pearson residual after GLM fitting shows the necessity of industry-specific hidden factors. We can incorporate observed common factors into our model as in previous studies, but will still need to decide on how to incorporated hidden factor into the model. The following sections discuss this issue.

2.4 Properties of the Model: Correlation Structure

In this section, a model for capturing correlation structures for within- and between-industry factors is introduced. To choose the type of time-series dependence and cross-sectional dependence of industry-specific hidden factors, a model selection procedure is conducted for three possible models of industry-specific hidden factors: Independent AR, SUR, and Multivariate AR models.

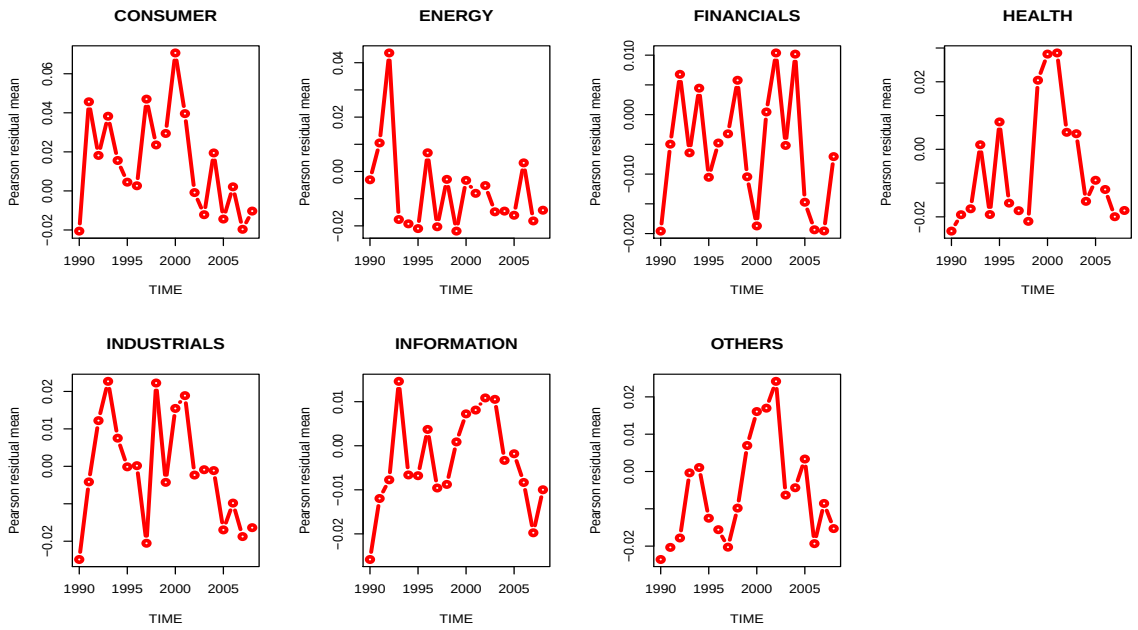

Figure 2.3: **Mean of Pearson Residuals.** Here GLM is fit with observed factors and the average of Pearson residuals is calculated in each industry in each year.

2.4.1 Model for Industry Hidden Effects

Even though each industry factor is modeled separately, it is hard to assume that industry factors are independent in terms of both time-series and cross-sectional aspects. We thus assume a Multivariate AR(1) type model for hidden industry effects, as shown in equation (2.3) and (2.4). In this setting, the matrix type K determines the correlation structure of industry effects with its own previous value and the previous values of other industries. The matrix type Σ determines the correlation structure between different industry effects at the same time.

In order to choose the matrix types K and Σ, we take the mean of the Pearson residuals after fitting the GLM as described in equation (2.6) for each month within each industry. With these mean of Pearson residual paths in each industry, we fit the three different models, which assume different matrix types K and Σ. The first model assumes diagonal matrix K and Σ, which is independent AR(1). The second model assumes diagonal K matrix and non-diagonal Σ matrix, which is the seemingly unrelated regression (SUR) model. The third model assumes non-diagonal matrix for both K and Σ, which is the multivariate AR(1) model. The Akaike information criterion

Table 2.4: **Model Selection Process.** This table shows the AICs and BICs for three possible models for hidden industry-specific factor: Independent AR (1), SUR and Full Multivariate AR (1) models.

Model	Independent AR	SUR	Full Multivariate AR
K	Diagonal	Diagonal	Non-Diagonal
Σ	Diagonal	Non-Diagonal	Non-Diagonal
AIC	-5533.953	-5676.041	-5416.18
BIC	-5342.156	-5556.168	-5388.719

(AIC) and Bayesian information criterion (BIC) are calculated for the three possible structures of Σ and K. Table 2.4 summarize the results.

According to the results, we assume a diagonal matrix K and a non-diagonal matrix for Σ. In other words, the j^{th} industry effect at time t, denoted by $\gamma_{j,t}$ is affected by its own previous value $\gamma_{j,t-1}$ and other industry effects only at the same time t.

Threrfore, our model for industry effect is

$$\begin{aligned} Y_{j,t} &= k_j Y_{j,t-\Delta t} + \varepsilon_{j,t},\ for\ all\ j = 1,...J, && (2.8)\\ Var(\varepsilon_{j,t}) &= \sigma_{j,j} && (2.9)\\ Cov(Y_{j,t}, Y_{l,t}) &= \sigma_{j,l} \neq 0\ where,\ j \neq l. && (2.10) \end{aligned}$$

2.4.2 Within- and Between-Industry Correlation

The merit of our model compared to previous models is that by adopting γ_{jt} as a random effect, we can define the within-industry and between-industry default correlation. By the law of total covariance, the covariance between the two default indicators $D_{i[j],t}$ and $D_{m[l],t}$ for firms i and m can be expressed as follows:

$$\begin{aligned} &Cov(D_{i[j],t}, D_{m[l],t})\\ &= Cov\left(E(D_{i[j],t}|\lambda_{i[j],t}), E(D_{m[l]t}|\lambda_{m[l],t})\right) + E\left(Cov(D_{i[j],t}, D_{m[l],t}|\lambda_{i[j],t}, \lambda_{m[l],t})\right)\\ &= Cov(\lambda_{i[j],t}, \lambda_{m[l],t})(\Delta t)^2 + 0 \qquad (2.11) \end{aligned}$$

The second part of equation (2.11) is equal to 0 because of the "doubly-stochastic" assumption in the intensity model, which means that, given the path of the state vector, the default times of different firms are conditionally independent. As a result, after controlling all observed common factors $U_{i,t}$ and X_t, default events are correlated because of correlated random industry effects $Y_{j,t}$ as follows:

1. Within-industry default correlation (after controlling all observed common effects) is (here firm i and m belong to industry j)

$$\begin{aligned} Cov(D_{i[j],t}, D_{m[j],t} | X_{1:t}, U_{i,1:t}, U_{m,1:t}, Y_{j,1:t-\Delta t}) &\propto Var(\exp(Y_{j,t}) | Y_{j,t-\Delta t}) \\ &\propto Var(Y_{j,t} | Y_{j,t-\Delta t}) = \sigma_{j,j}. \end{aligned} \tag{2.12}$$

2. Between-industry default correlation(after controlling all observed common effects) is (here firm i is in industry j and m is in industry l)

$$\begin{aligned} &Cov(D_{i[j],t}, D_{m[l],t} | X_{1:t}, U_{i,1:t}, U_{m,1:t}, Y_{j,1:t-\Delta t}, Y_{l,1:t-\Delta t}) \\ &\propto Cov(\exp(Y_{j,t}), \exp(Y_{l,t}) | Y_{j,t-\Delta t}, Y_{l,t-\Delta t}) \\ &\propto Cov(Y_{j,t}, Y_{l,t} | Y_{j,t-\Delta t}, Y_{l,t-\Delta t}) \\ &= \sigma_{j,l}. \end{aligned} \tag{2.13}$$

It is impossible to get exact between- and within- industries default correlation coefficients through this method, but we can still compare their relative scale through estimated Σ which is the conditional covariance matrix of random industry factors. To show the effects of the variance and covariance of Y, we conduct a simple simulation. The monthly data of 70 firms for 60 months is generated. In line with our empirical data, seven industry categories are set and 10 firms are assigned in each category. We first simulate Y_t with varying values of the diagonal elements of Σ: $0.1, 1.3$ and 2.5. All the off-diagonal elements of Σ are fixed to 0. We then generate the three hypothetical data sets for all three corresponding diagonal matrices of Σ and simulated Y_t vectors. Other than industry-specific random effects $Y_{j,t}$, we control parameters in equation (2.2)

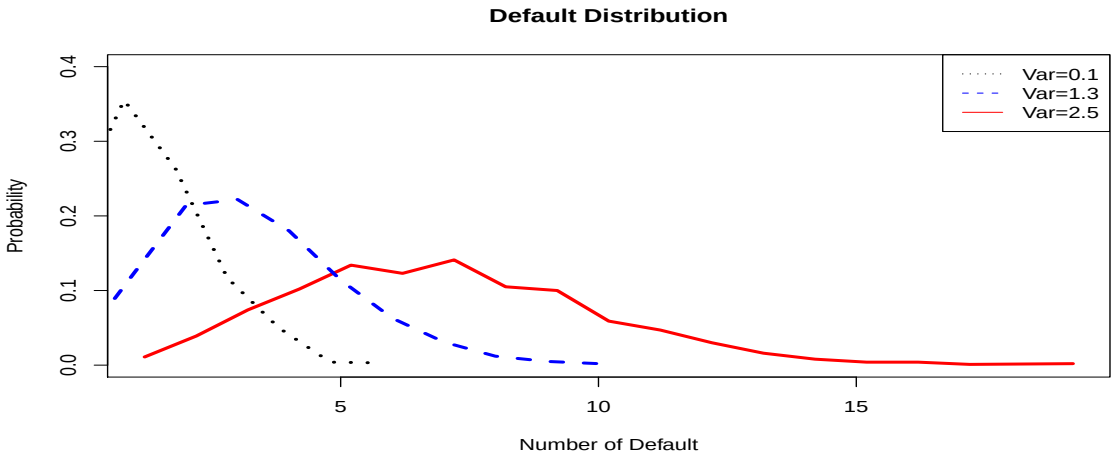

Figure 2.4: **Simulated Default Distributions According to Different Variance Values.** We fix the parameters as $a_0 = -8$ and $a_1 = a_2 = 0$. We fix all correlation values (off-diagonal elements of Σ) to 0.

with $a_0 = -8$ and $a_1' = a_2' = (0,0)$. Finally, we simulate default time for each hypothetical data set. By repeating these simulation steps 1000 times, we can get the default distribution for each different diagonal element in Σ.

Figure 2.4 shows the simulated default distributions for each different diagonal element in Σ. We find that, as the variance of Y (diagonal element of Σ) increases, the default distribution shifts to the right and has a longer tail.

Figure 2.5 shows the simulated default distributions for three different off-diagonal values in Σ. We set the diagonal value of Σ to 2.5 in all three cases, thus off-diagonal values we tested are $0 * 2.5, 0.5 * 2.5$ and $0.99 * 2.5$, respectively. As the correlation of Y increases, the default distribution has a longer right tail.

Even though it is hard to extract the exact default correlation coefficients of between- and within-industries through our model, we can still obtain the relative size of each. We can infer that if the default probabilities are the same, then the industry with a large variance of Y has a larger within-industry correlation and more defaults occur than in other industries; pairs of industries with a large covariance between Y also have a larger between-industry correlation and more defaults occur than in other pairs.

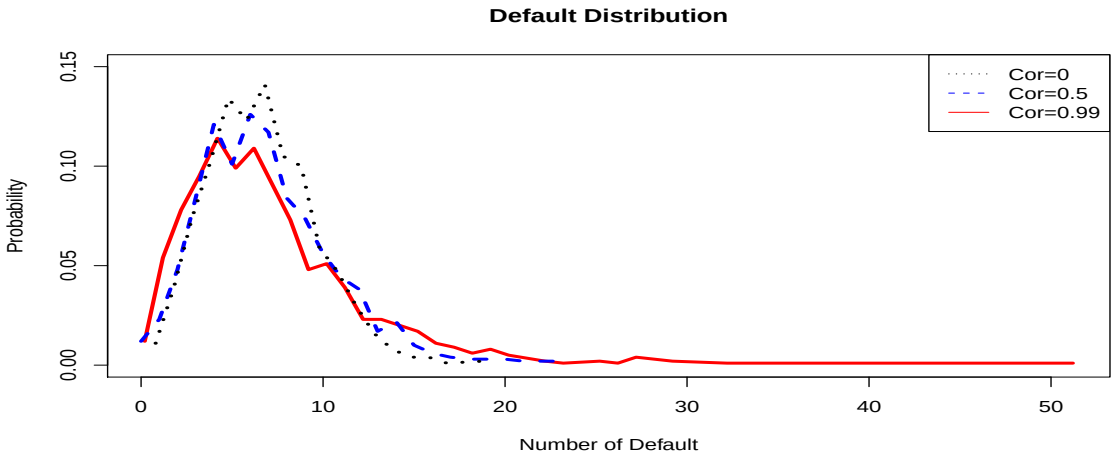

Figure 2.5: **Simulated Default Distributions According to Different Correlation Values.** We fix the parameters as $a_0 = -8$ and $a_1 = a_2 = 0$. We fix all variance values (diagonal elements of Σ) to 2.5.

So far the discussion has centered on our model and how to measure and compare between- and within- industry correlations using this model. In the next section, we introduce the Bayesian analysis method adopted in this essay to estimate model parameters.

2.5 Estimation

As our model becomes more sophisticated in order to incorporated industry-specific correlation, it takes on added complexity in estimation procedure. The parameters of interest are the fixed effect coefficients vector $a = (a_0, a_1, a_2)$, the AR coefficients matrix of random effects K, and the covariance matrix of the random effect Σ. The hidden industry effect $\vec{Y}_t = (Y_{1,t}, ...Y_{J,t})$ for $t = 1, ...T$ also needs to be estimated, where T is the last time in the data set and J is the total number of industries. The complete likelihood function $L(a, K, \Sigma | Y, X, U, D)$ is as shown below:

$$L(a,K,\Sigma|Y,X,U,D) = \prod_{i=1}^{n} \left(exp(-\sum_{t=t_i}^{T_i} \lambda_{it}\Delta t) \prod_{t=t_i}^{T_i} [D_{it}\lambda_{it}\Delta t + (1 - D_{it})] \right) \prod_{t=1}^{T} p(Y_t|Y_{t-1}, K, \Sigma), \tag{2.14}$$

where λ_{it} is a function of X, U, Y and a as shown in equation (2.2), $p(Y_t|Y_{t-1}, K, \Sigma)$ is the normal density function as shown in equations (2.3) and (2.4), X is a matrix of observed common factor

vectors, U is a matrix of observed firm specific factor vectors, I is the total number of firms, and $D_t = (D_{1,t}, ...D_{I,t})$ is a vector of default indicators. T_i refers to the last observation time of company i, which could be the time of default or time of exit for other reasons. We take the first appearance time t_i to be deterministic.

In order to estimate not only a, K and Σ, but also $Y = (Y_1, ...Y_T)$, we use the Monte Carlo EM (MCEM) algorithm. The MCEM algorithm can be adopted if the E-step is analytically difficult to compute, as in this essay. The idea behind the MCEM algorithm is that the E-step can be modified by using posterior sampling (Wei and Tanner (1990)). As a first step to implementing the MCEM algorithm, we simulate the hidden factor with a Metropolis-Hasting (MH) algorithm. The individual hidden factor value $Y_{j,t}$ is the (j,t) element of Y, where $j = 1, ...J$ refers to industry, and $t = 1, ...T$ refers to time. The target distribution for individual $Y_{j,t}$ is

$$P(Y_{j,t}|Others) \propto \prod_{i=1}^{n_j} (exp(-\lambda_{it}\Delta t)[D_{it}\lambda_{it}\Delta t + (1 - D_{it})])\, p(Y_{j,t}|Y_{j,t-1}, K, \Sigma) p(Y_{j,t+1}|Y_{j,t}, K, \Sigma), \tag{2.15}$$

where n_j is total number of firms in the j^{th} industry. As a proposal distribution for the MH algorithm, we use the normal distribution with $s1$ tuning parameters controlling the step size as below:

$$p(Y_{j,t}^m | Y_{j,t}^{m-1}) \sim Normal(Y_{j,t}^{m-1}, s1). \tag{2.16}$$

In each E-step, we increase the posterior sample size with iteration with rate 5 up until size 100.

After Monte Carlo sampling for the hidden factor, we then proceed to the M-step with a simulated Y. The M-step is the procedure to obtain parameters that maximize the likelihood function in equation (2.14). We can separate maximization for a and (K, Σ) because the complete likelihood function shown in equation (2.14) is simply a product of $L(a|Y, X, U, D)$ and $L(K, \Sigma|Y, X, U, D)$. The first maximization is for a, so we need to find $\hat{a}$, which maximizes the likelihood function of a, and which is given as $L(a|Y, X, U, D) = \prod_{i=1}^{n} \left(exp(-\sum_{t=t_i}^{T_i} \lambda_{it}\Delta t) \prod_{t=t_i}^{T_i} [D_{it}\lambda_{it}\Delta t + (1 - D_{it})] \right)$. This maximization is equivalent to fitting a GLM, and can therefore be accomplished using the standard iteratively reweighted least squares (IRLS) algorithm (Mcculloch and Searle (2001); Zipunnikov and Booth (2006)). The second maximization is for K, Σ. $\hat{K}, \hat{\Sigma}$ are determined where

Table 2.5: **Estimated Fixed Effects.** Maximum Likelihood Estimates (MLE) of fixed effects of our industry-specific hidden factor model. Intercept terms is fixed as same as the *without hidden factor model* because of its unique identification ($a_0 = -0.012$ which is shown in Table 2.7).

	Estimate	Std. Error	z value	$Pr(>\|z\|)$
DTD($a_{1,1}$)	-2.242	0.112	-19.943	$< 2e-16$ *
lag12($a_{1,2}$)	-0.001	0.029	-0.036	0.971
sp500($a_{2,1}$)	1.243	0.305	4.080	$4.5e-05$ *
Tb3($a_{2,2}$)	0.006	0.022	0.260	0.795

$L(K,\Sigma|Y,X,U,D) = \prod_{t=1}^{T} p(Y_t|Y_{t-1},K,\Sigma)$ is maximized.

2.6 Empirical Results

2.6.1 Observed Factor Effect: $\hat{a}$

Table 2.5 shows the estimated fixed effect $\hat{a}$. As an initial configuration for a, we use coefficients estimated without the hidden factor. We fix the intercept term a_0 with the value estimated without the hidden factor because of its unique identification. The mean level at time t in the j^{th} industry is $a_0 + Y_{j,t}$, so without fixing a_0, $Y_{j,t}$ is not uniquely determined.

Table 2.5 shows the important roles for both firm-specific and macroeconomic variables. The distance to default is the most significant variable. For example, a negative shock to distance to default by one standard deviation increases the default intensity by roughly $exp(2.242) - 1 \approx 841\%$. The firms trailing one-year stock return have the smallest effect on the default intensity. This covariate are introduced in Shumway (2001) as a proxy for firm-specific information that is not captured by distance to default. However, it cannot perform an important role for our data. Among macroeconomic factors, the S&P500 return is the most significant variable. The signs of S&P500 to a firm's default intensity are positive, which was an unexpected sign by univariate reasoning. As Duffie, Eckner, Horel, and Saita (2009) mentioned, this is because of the model dominance of distance to default. Normally, after a boom year, a firm's distance to default tends to overstate its financial health.

Table 2.6: **Estimated Correlated Structure of Industry-Specific Random Effects.** MLE of Σ and K of our industry-specific hidden factor model.

Industry	Cons	Energy	Finan	Health	Indus	Info	Oth
AR coefficients	0.691	0.923	0.853	0.883	0.800	0.785	0.873
Variance	1.053	0.167	0.303	0.375	0.504	0.569	0.488
Correlation	Cons	Energy	Finan	Health	Indus	Info	Oth
Consumer	1.000	0.784	0.811	0.886	0.887	0.811	0.902
Energy	0.784	1.000	0.730	0.734	0.730	0.696	0.789
Financials	0.811	0.730	1.000	0.853	0.898	0.930	0.909
Health	0.886	0.734	0.853	1.000	0.925	0.922	0.952
Industrials	0.887	0.730	0.898	0.925	1.000	0.936	0.918
Info-Tech	0.811	0.696	0.930	0.922	0.936	1.000	0.924
Others	0.902	0.789	0.909	0.952	0.918	0.924	1.000

2.6.2 Estimated Random Effect: $\hat{K}$, $\hat{\Sigma}$

As shown in subsection 2.4.2, under our industry-specific hidden factor model, the covariance and variance terms of the random effect Y measure the between-industry correlation and within-industry correlation, respectively. Table 2.6 shows the estimated variance and correlation matrix of an industry-specific random effect.

While we cannot find an exact between- and within- industries correlation through this result, we can still compare the relative scale of correlations for the firms with same observed factor values. If all other conditions(firm-specific variable values, observed common factor values) are same, the consumer discretionary industry has the largest within-industry correlation, and the energy industry has the smallest within-industry correlation. The magnitude of the AR coefficients, however, are in the opposite order. The energy industry has the largest AR coefficient, and the consumer discretionary industry has the smallest. Except for these two, there are no sizeable gaps in terms of AR coefficients.

The information technology industry has relatively larger between correlations. As expected, the energy industry, which shows most unique pattern, has smaller between correlations, in general. As we expected from before (the Figures 2.2 and 2.3), the financials industry correlated most with the information-technology industry and the consumer discretionary industry correlated most

with the industrials industry. The highest correlations are found among the health care, industrials, and information technology industries.

2.7 Model Comparison

In this section, as a comparison study, we investigate how our industry-specific model is improved compared to previous correlated default models. The following three models are compared:

1. *Without hidden factor model*:

$$\lambda_{i,t} = \exp\{a_0 + a_1 U_{i,t} + a_2 X_t\} \tag{2.17}$$

2. *Common hidden factor model*:

$$\lambda_{i,t} = \exp(a_0 + a_1 U_{i,t} + a_2 X_t + Y_t) \tag{2.18}$$

3. *Industry-specific hidden factor model*:

$$\lambda_{i[j],t} = \exp\{a_0 + a_1 U_{i,t} + a_2 X_t + Y_{j,t}\} \tag{2.19}$$

2.7.1 Comparison of Estimated Observed Factor Effect and Hidden Factor

In this subsection, we show the difference in estimated observed factor effects and hidden factor processes. For comparison with our *industry-specific hidden factor model*, we give the estimated $\hat{\beta}$ in the *without hidden factor model* and the *common hidden factor model* in Table 2.7. The estimated fixed effect of the *industry-specific hidden factor model* is given in Table 2.5.

There are no significant changes in the effects of firm-specific variables. Similar to *the industry specific model*, the most significant factor is distance to default, and the firm's trailing one-year stock return is the least significant variable in all three models. The sign, magnitudes, and statistical significance of distance to default variable are similar in all three models. However, after adding the hidden factors, there is a change in the effects of macroeconomic variables. Af-

Table 2.7: **Estimated Fixed Effect with Previous Models.** These are MLEs of fixed effect of without hidden factor model and common hidden factor model. Intercept term of the *common hidden factor model* is fixed as same as the *without hidden factor model* because of its unique identification.

Without Hidden Factor Model

	Estimate	Std. Error	z value	$Pr(>\|z\|)$
(Intercept)	-3.012	0.145	-37.849	$< 2e-16$ *
DTD	-2.303	0.143	-16.142	$< 2e-16$ *
lag12	0.002	0.027	0.090	0.929
sp500	0.190	0.311	0.612	0.541
Tb3	0.146	0.029	4.929	$8.28e-07$ *

Common Hidden Factor Model

	Estimate	Std. Error	z value	$Pr(>\|z\|)$
DTD	-2.280	0.111	-20.584	$< 2e-16$ *
lag12	0.0007	0.026	0.026	0.979
sp500	-0.070	0.307	-0.227	0.820
Tb3	0.162	0.021	7.557	$4.13e-14$ *

ter adding the common hidden factor, the effect of the T-bill rate slightly increases; the effect of S&P500 decreases and its sign become negative. However, there is no change in terms of their significance.

In contrast, after adding the industry-specific hidden factor, there are large changes in the opposite direction of both in terms of their magnitude and significance. The effect of a T-bill rate decreases and becomes insignificant; the effect of S&P500 increases and becomes significant. Estimated hidden factor processes account for the reason for differing changes in these instances. Figure 2.6 shows the estimated hidden factor processes in the *common hidden factor model* and *industry-specific hidden factor model*. Here, estimates refer to the posterior means of sampled hidden paths at parameter converging iteration. In the *common hidden factor model*, except during the early 1990s and after 2005, hidden factor values fluctuate near 0, so adding the hidden factor does not affect the estimate of observed common factor effects. In contrast, industry-specific hidden factor values vary a great deal, not only within each industry, but also between industries.

The fluctuations of the consumer-discretionary, financials, industrials and information-technology industries' hidden factors are similar to the common hidden factor movement. In contrast, the

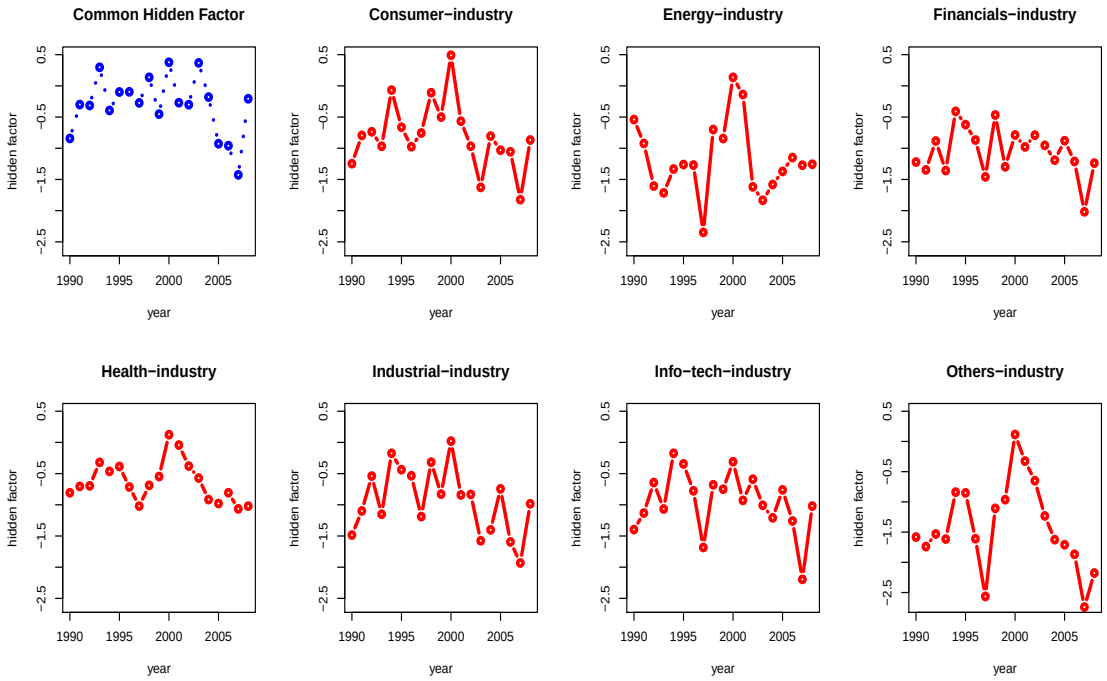

Figure 2.6: **Estimated Hidden Factor Process in Common Hidden Factor Model and Industry-Specific Hidden Factor Model**. The first plot (blue dotted line) is the estimated hidden factor of the common hidden factor model. The remaining seven plots (solid red lines) are estimated industry-specific hidden factors based on the *industry-specific hidden factor model*. The estimated hidden factor value is the mean of posterior mean of the sampled hidden factor process in each year at parameter converging iteration .

energy, health care, and "others" industries reveal distinguished patterns in estimated common hidden factor. There is a significant drop in 1997 and an increase after that until 2000; there is no significant movement between 2006 and 2008. The hidden process patterns of these industries have relatively high correlation with T-bill rates and with the S&P500 time series when compared to other as follows: The correlation coefficient between T-bill rate and the posterior mean of the energy industry hidden factor process is about 0.41, while correlation coefficients between T-bill rate and the posterior mean of the hidden factor processes of the consumer discretionary, financials, industrial and information technology industries are in the range from -0.026 to 0.13. Similarly, the correlation coefficients with S&P500 are all near -0.2 in the energy, health care, and others industries while in discretionary, financials, industrials and information technology industries, they are less than -0.1, mostly around -0.01.

Throughout this comparison between the hidden factor processes in the *industry-specific* and *common hidden factor models*, we can infer that the effect of the T-bill rate is over-valued without the industry-specific hidden factor because the estimated effect of the T-bill rate is dominated by companies in the energy, health care, and "others" industries. Even though we include the common hidden factor, the unique patterns emerging from these industries cannot be captured because the common hidden factor process is determined by a majority of other firms having a similar hidden factor process, and their unique patterns are averaged out by assuming one common hidden factor for entire industries. Therefore the over-valued T-bill rate effect cannot be adjusted only with the common hidden factor when we need the industry-specific factor.

2.7.2 Comparison of Estimated Default Distribution

In this subsection, our comparative study shows how the *industry-specific hidden factor model* performs best for predicting the number of defaults in each year. In order to derive the default distributions, for the *common hidden factor model* and the *industry-specific hidden factor model*, we sample hidden factor processes with estimated parameters described in Table 2.5 and Table 2.7. We iterate 100 times the simulation of hidden factor processes, default intensity processes, default times, and total number of defaults in each industry and for each year. For the *without hidden factor model*, we simply iteration 100 times the simulations of default intensity process and default times with estimated parameters.

$RND_{j,t}$ denotes the real number of defaults in j^{th} industry in year t. $SND_{j,t}^{(l)}$ denotes the simulated number of defaults in j^{th} industry in year t at l^{th} iteration. Then q_{jt}, which is the posterior cumulative distribution of realized number of defaults in each year and in each industry, is defined as below:

$$q_{jt} = \frac{1}{100} \sum_{l=1}^{l=100} I_{jt}^{(l)}, \tag{2.20}$$

where

$$\begin{aligned} I_{jt}^{(l)} &= 1 \ if \ SND_{j,t}^{(l)} \leq RND_{j,t} \\ &= 0 \ if \ SND_{j,t}^{(l)} > RND_{j,t}. \end{aligned} \tag{2.21}$$

Figure 2.7 shows q_{jt} for all industries and years.

In the *without hidden factor model*, most of q_{jt} cluster near 0 or 1. For the period around 2000, the *without hidden factor model* underestimates the number of defaults (default probability), and from 2005 to 2008, it overestimates. Moreover, these over- or under-estimated periods are clustered. For example, from 1998 to 2002, the realized number of defaults lies above the estimated 3^{rd} quantile of the portfolio default distribution in entire industries and all individual industries. In contrast, from 2005 to 2008, the realized number of defaults lies below the estimated 1^{st} quantile of the portfolio default distribution in entire industries and all individual industries.

Incorporating hidden factors help to resolve these over- and under-estimation problems for entire industries. However, in the analysis of individual industries, the *common hidden factor model* still has over- or under-estimation problems. For the period around 2000, the *common hidden factor model* underestimates the default probability in the consumer discretionary and health care industries, but in the energy and financial industries, it overestimates the default probabilities. The over- and under-estimation periods are not the same in different industries. In particular, the consumer discretionary and energy industries have significantly different patterns.

The industry-specific hidden factor model gives more centered $q_{j,t}$ and is distributed relatively evenly in the unit interval for the entire industry and for all individual industries. This indicates a more accurate assessment of credit risk in all cases than the other two models. Figure 2.8 and Figure 2.9 show this result in clearer terms.

Two periods, 1999-2001 and 2005-2007, are selected. The first period represents high default years, and the second period represents low default years. Figure 2.8 and Figure 2.9 depict the posterior distribution of the number of defaults under the three models, with the realized number of defaults during two selected periods, respectively.

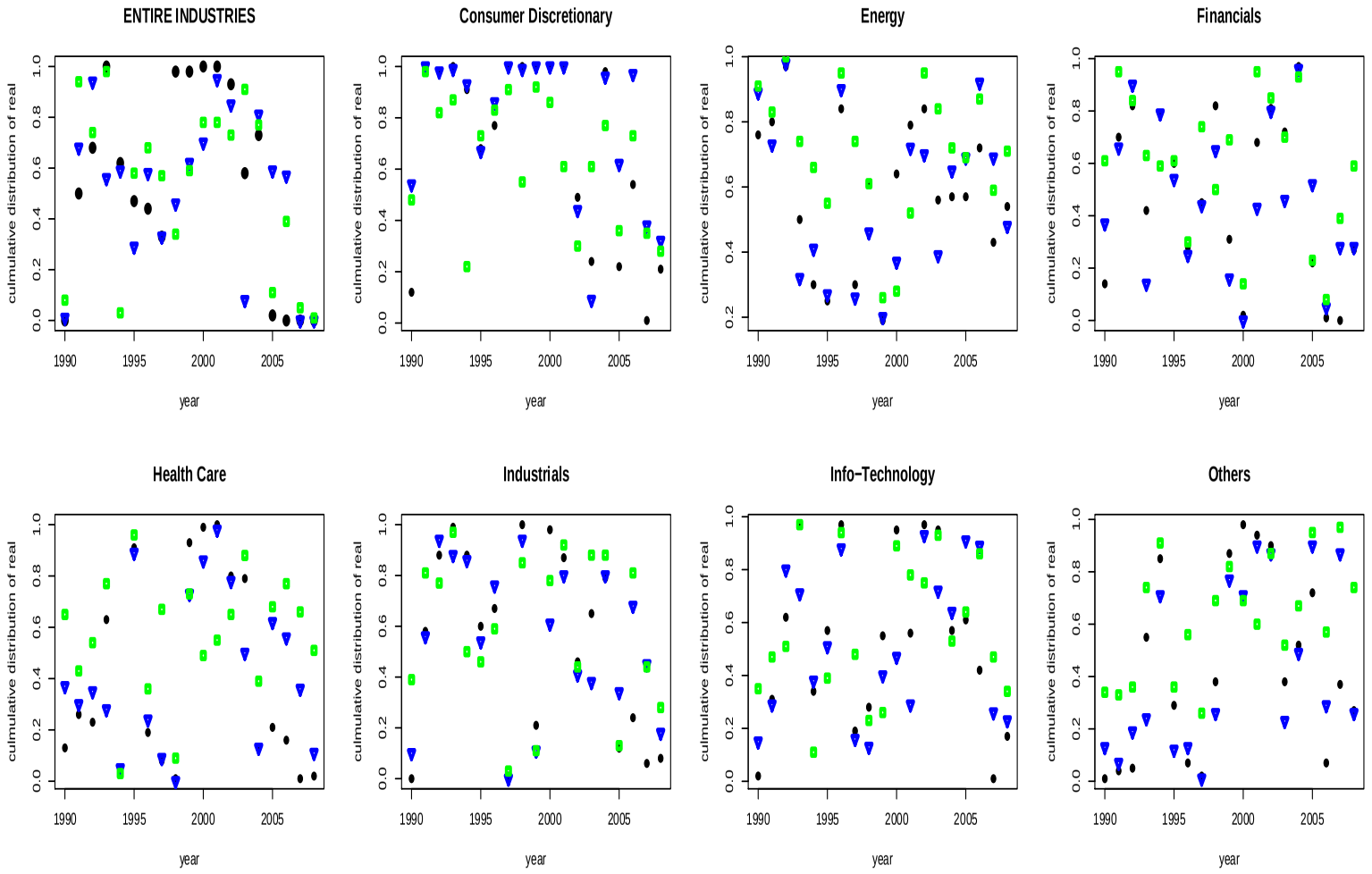

Figure 2.7: **Posterior Cumulative Distribution of Realized Number of Defaults.** Defaults distribution are simulated under three models: without hidden factor model (black circle), common frailty model (blue triangle) and industry-specific hidden factor model (green rectangular).

During the low default years (from 2005 to 2007) shown in Figure 2.9, there is no marked difference between the *common hidden factor model* and the *industry specific hidden factor model*. The hidden factor in both models takes a major role in adjusting over-estimated default probability. However, during the high default years (from 1999 to 2001) shown in Figure 2.8, we see a need for the industry-specific hidden factor. While there is no sizable difference between *common hidden factor* and *industry-specific hidden factor models* in terms of entire industries, we see the difference at the level of individual industries.

During the high default years, the common hidden factor model moves default distribution to the right at the same rate for all industries. As a result, the number of defaults in the consumer discretionary and health care industries are under-estimated because these industries suffer the relatively large increase of defaults during this period. In contrast, the defaults distribution is over-estimated in the financials industry when the *common factor model* is used.

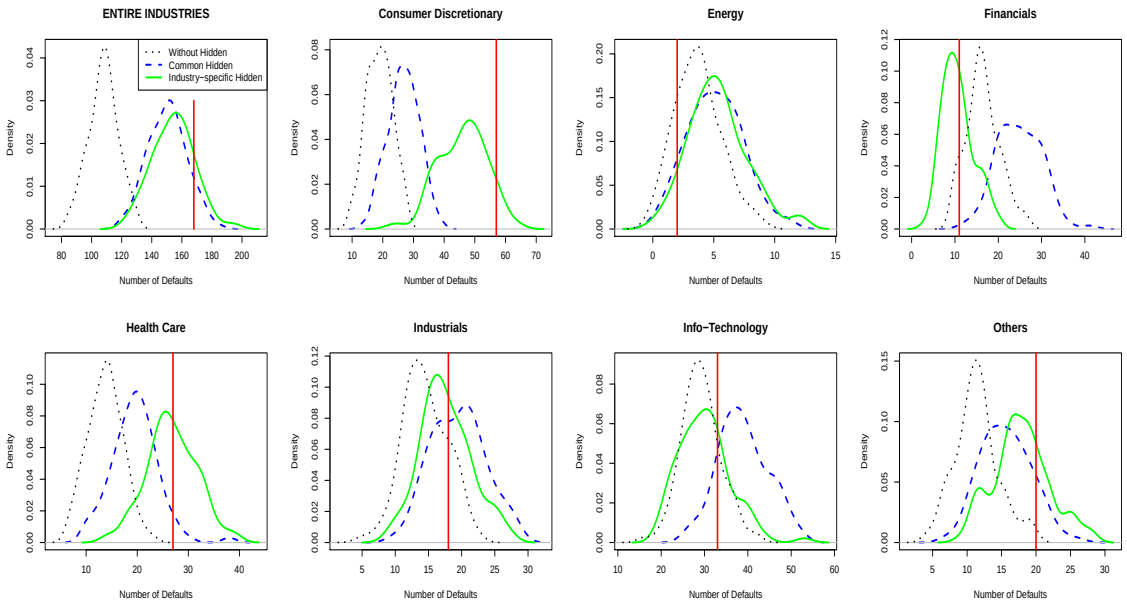

Figure 2.8: **Estimated Defaults Distributions between 1999 and 2001.** The without hidden factor model is marked with a black dotted line; the common hidden factor model is marked with a blue dashed line; the industry-specific hidden factor model is marked with a green solid line. The red vertical line shows realized number of defaults.

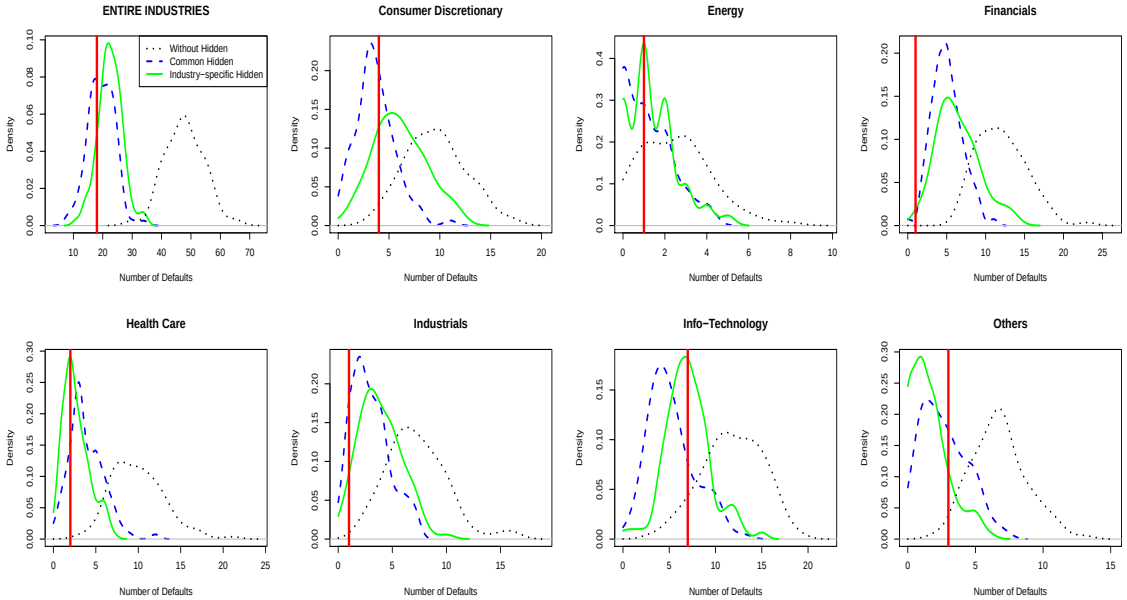

Figure 2.9: **Estimated Defaults Distributions between 2005 and 2007.** The without hidden factor model is marked with a black dotted line; the common hidden factor model is marked with a blue dashed line; the industry-specific hidden factor model is marked with a green solid line. The red vertical line shows realized number of defaults.

2.8 Conclusion

In this essay, we extended previous studies by adopting industry-specific hidden factors as random effects. By introducing industry-specific hidden factors and assuming that they are random effects, we arrived at a relative scale of within-industry and between-industry correlations across a broad cross-section of industries. Under the condition of unknown hidden paths and analytically-difficult likelihood functions, the MCEM algorithm was used to estimate parameters and the random effect process.

Our estimates were derived from large data sets both in terms of cross-sectional and time-series aspects. We collected monthly data during 19 years for the 13,682 U.S. public firms, which is $1,413,314$ firm-months of data, from January 1990 to December 2008. Therefore, our estimated results can be applied for future prediction with more accuracy.

Instead of fixed effects for the time-varying common factor, we applied an industry-specific hidden factor and measured within- and between-industry correlations. We compared the relative

scale of within- and between-industries correlations. As expected, the energy industry, which shows a unique default rate time series, has the smallest within- and between-correlations.

In order to prove the necessity of the industry-specific hidden factors, we compared three models: *without hidden factor model*, Duffie's *common hidden factor model*, and our i*ndustry-specific hidden factor model*. There is no significant gap between the *without hidden factor model* and the *common hidden factor model* in estimated parameters, but adding the industry hidden factor makes a substantial difference in estimating the observed common factor effect. We also showed that the effect of T-bill rate is over-valued without the industry-specific hidden factors, because the T-bill rate has significant impact only in industries in the energy, health care, and others categories. In terms of the default distribution prediction, without industry-specific hidden factors, there are still over- and under- estimation of default probability in some industries during some years.

The model built in this essay has a potential contribution in prediction. By building a time-series model for industry-specific hidden factor, an industry-specific prediction can be made. Our results also have important implications for the more sophisticated risk management strategy. For example, if investors wish to diverse their risk more, then they could compose their investment portfolio with the bonds in less correlated industries. Moreover, when large default events happen, such as Enron in 2001, WorldCom in 2002 and Lehman Brothers in 2008, investors can expect the effect of defaults to other firms' defaults more precisely in the industry-specific sense.

The model described and analyzed in this essay is only about the default event, thus it is useful in defaultable bond price modeling or prediction. However, there are many other reasons for firms' failures in the market. As shown in Table 2.2, default is an extreme event and explains few of the reasons a firm exits the market. In the financial market especially, other reasons such as merges and acquisitions are largely responsible for the majority of firms' failures, and that is why estimated within industry correlation in the financial market is not as large as we expected at first. However, our approach can be easily extended to the analysis on other failure reasons with similar structure. Combining the results on all failure reasons of the firms might give better understanding on the impact of certain firm's or industry's collapse.

3. PARALLEL INTENSITY AND STRUCTURAL MODELS FOR CORRELATED DEFAULTS

3.1 Introduction

Under the two main credit risk models, structural and intensity models, default dependence is modeled with asset correlation and default intensity correlation, respectively. Based on each of these model settings, techniques to deal with correlated defaults have been independently developed. In this essay, by combining the merits of several default correlation studies, we propose a common model framework for multi-firm defaults under the structural and intensity credit risk model settings and estimate the model using multiple securities.

The key ideas behind the combined model structure are the following. First, from structural model studies, we adopt the state-space model approach. We assume that each firm's unobserved state process follows certain stochastic processes and these determine the default time and other financial market values. Under the structural model, the state variable is the asset process, and under the intensity model, the state variable is the intensity process. Second, to capture the correlation structure, the common factor method is adopted from studies of the intensity model. Based on the common factor method, correlation of defaults is explained by the state processes, which share time-varying common factors. Both observed and hidden common factors are adopted in this essay, as in Duffie, Eckner, Horel, and Saita (2009).

With such a combined model structure, the main contribution of this essay is that we combine the advantages of both typical intensity and structural models. First, by adopting the common factor method, it is possible to consider default correlation among multi-firms more easily than in typical structural models. Second, by assuming specific stochastic processes for each state process, empirical analysis does not need to rely on extensively long-term or massive amounts

of firm data, which also include the parts that are irrelevant to our needs. In order to adopt the factor method in typical intensity models, a large data set (both in terms of number of firms and time points) is necessary to ascertain that the data set contains suitably many defaults to carry out a statistical analysis. This is because a default event is very rare. In fact, Duffie, Eckner, Horel, and Saita (2009) used the monthly data from 2,793 companies over a 25-year period, which is more than 400,000 firm-months data. Third, state-space model approaches also make it possible to easily pool different securities prices, and utilize them together in estimation because the capital structure of a firm can be evaluated as a whole in one consistent framework.

In addition to adopting a parallel structure within two different credit risk models, our study has the following distinctions from previous default correlation studies. First, instead of completely pooling all the firms into a single model, by adopting random coefficients, two extreme assumptions are compromised: strictly different factor effects and equal factor effects across companies. We assign a probability distribution on the coefficients and estimate not only the coefficients themselves, but also the parameters (hyperparameters) governing their probability distribution with data. Such hyperparameters of random coefficients make it possible to measure the degree to which random coefficients are connected and the general effect of corresponding common factors. Furthermore, comparison among different factor effects within structural or intensity models, and comparison of certain factor effects between structural and intensity models, are all possible through hyperparameter comparison.

Second, the CDS spreads of the underlying firms are additionally used for estimation. For the intensity model, because it is difficult to apply stock price information, we depend only on the CDS spread information, as in standard practice. However, in the structural model, we first use stock prices information (equity prices)[1]as is standard practice, and then add the CDS spread information in the estimation process. Unlike previous uses of CDS prices in the structural model, the CDS prices and equity price are simultaneously used to estimate the structural model parameters and all hidden processes. We have already shown in our first essay that the additional securities information enhances estimation accuracy and efficiency. However, the option prices we used in the first essay are already dependent on the stock prices; thus there is a limitation in terms of new

information. On the other hand, because the market price of CDS should reflect market assessment of credit events such as default, CDS spreads are known to serve as a good source of information in estimating the credit risk of firms and firm default correlations.

Our empirical results derived from daily one-year 2009 data and they support the necessity of hidden factor and relaxation of assumptions we made in this essay. Results depicting that each factor coefficient is widely spread across different firms show the necessity of relaxation of the equal factor coefficients assumption. We also show that, in both structural and intensity models, the hidden factor can capture the common movement in the underlying state process not captured by the observed common factors. In terms of the necessity of incorporating CDS information, empirical study indicates that adopting CDS prices makes it possible to estimate the default intensity process and to enhance the precision of the estimation in the asset value process to a greater extent than when using only the stock market information.

The third distinction of our study from previous work is in market noise. The market noise in the CDS and stock markets are all considered in this essay. Market noise makes it possible to combine stock and CDS market information in estimation, and we can incorporate both model mis-specification error and market il-liquidity. In our intensity model, we also allow CDS market noise, while the typical intensity model assumes there is a one-to-one relationship between 5-year CDS price and hidden intensity value.

As an estimation method, the Bayesian method is adopted, which estimates together all unknown state processes, the hidden factor process and parameters. Gibbs sampling and the Metropolis-Hasting algorithm are the basic methods used in this essay. However, these methods are used with the simultaneous jumping (generalized Gibbs) technique to improve chain convergence problems (Kou, Xie, and Liu (2005) and Liu (1996)).

Before we conclude the analysis, in order to perform model validation checks and show potential uses of this essay, we conduct CDX (CDS Index) tranche prices prediction. By virtue of the Bayesian method, the posterior distribution of CDX tranche prices can be driven, even though it is

[1] We will keep using *stock* and *equity* without any distinction. They are different terminologies in finance, but we calculate equity as a product of stock price and number of outstanding. Equity referred in this essay is actually stock market information.

not used in the estimation procedure. By doing so, comparison between our estimated correlated structural model and intensity model is possible in the sense of predicting real financial market. Incorporating the correlation structure improved the prediction of CDX tranche market quotes, especially under the correlated structural model. However, in the intensity model, prediction did not work as well because all observed factors only marginally impact the intensity.

The rest of this essay is constructed as follows. Section 3.2 gives background knowledge of finance and previous research. Section 3.3 describes preliminary results related to the selection of common factors, including the hidden factor and random coefficients. In section 3.4, we discuss the common structure of our two correlated credit risk models determined by the preliminary study. Section 3.5 introduces our first default correlation model (correlated structural model), and section 3.6 presents the second model (correlated intensity model). Section 3.7 provides details of our Bayesian estimation procedure. In section 3.8, we investigate the performance of our Bayesian estimation method using simulated data. Section 3.9 is the introduction to our data set and adjustments of unknown values such as recovery rate and default barrier for model identification. Section 3.10 gives the empirical results. Section 3.11 present the CDX tranche prices prediction results based on estimated structural and intensity models. The conclusions are presented in section 3.12.

3.2 Background in Finance

3.2.1 Credit Risk Model and Default Correlation Studies

There are differences in defining the default event between two main streams in credit risk model research: structural model and intensity model. The structural model assumes an asset process follows geometric Brownian motion, and when the asset value of the firm falls below a certain threshold, the default occurs. Under this basic setting, there are variations in the model based on assumption involving the default time and default boundary. In this essay, we will adopt the Black-Cox structural model. Under the Black-Cox structural model, the default time τ is the

first time when asset value V_t hits the default boundary $L(t)$ as follows:

$$\tau = \inf\{t; V_t \leq L(t)\}. \tag{3.1}$$

In the structural model, standard practice suggests estimating asset values from stock price using the Merton model (Merton (1974)) and obtaining asset correlation. In a more theoretical sense, Zhou (2001) examined the structural model, which is extended to two firms whose assets are correlated, and found a closed form formula for the joint default probability of two issuers. More recently, in a manner similar to that of the factor method in typical intensity models, Hull, Predescu, and White (2010) provided a way to model correlation under the structural model. They divided the brownian motion governing asset W_{it} into a common component X_t and an idiosyncratic component W_{it}^{ε} and then measured the degree of sensitivity to the common factor with a_i as follows:

$$dW_{it} = a_i dX_t + \sqrt{1 - a_i^2} dW_{it}^{\varepsilon}. \tag{3.2}$$

With this a_i, they measured asset correlation.

Unlike structural models, defaults are here modeled as stochastic events whose arrival rates are governed by the given intensities. Let E_1 be the exponential random variable with mean 1; then under the intensity model, the default time τ is defined as below:

$$\tau = \inf\{t; \int_0^t \lambda_s ds \geq E_1\}, \tag{3.3}$$

where λ_s is default intensity.

The parameters governing intensities are typically assumed to depend on a set of market data. Therefore, under the intensity models, in order to handle correlation among a large number of firms, a method of imposing factor structure on default intensities method has been studied (Duffie and Garleanu (2001)). Duffie and Wang (2007) provided maximum likelihood estimators of term structures of conditional probabilities of corporate default by incorporating macroeconomic covariates in addition to firm specific variables. Recently, Duffie, Eckner, Horel, and Saita (2009)

applied the frailty model by adopting the hidden factor and adjusted the bias of default distribution estimation in Duffie and Wang (2007).

A common complexity in defaults analysis under intensity and structural models is due to the fact that the underlying variables, which are critical elements in each model-asset value and default intensity- are not observed. Therefore, research on this complexity derive the inference on unknown underlying values from the observed market prices. Under the structural model framework, because asset values are defined as a sum of bond and equity values, the usual estimation method is based on equity price to estimate the asset process and its parameters. In Chapter 1, such an estimation method is given with details under the Black-Cox structural model. In the intensity model, the usual estimation method adopts the CDS spreads as observations. In this essay, with consideration of the market noise, we apply both CDS and equity price to estimate the structural model and use CDS market information for estimation of the intensity model. Therefore, before discussing the model details and estimation method, as financial background, we introduce the CDS market.

3.2.2 CDS Market

As the CDS market continues to grow rapidly, there are a number of recent studies on the incorporation of CDS data into the credit risk model. Friewald (2009) adopted CDS data to estimate asset correlation and then compared this result with the estimation result using equity prices. Pan and Singleton (2008) estimated parameters governing default intensities with CDS prices of three countries, Korea, Turkey, and Mexico. Turnbull and Yang (2008) also used CDS prices to estimate jump intensities and parameters related to default barriers. After estimating model parameters in the intensity model and the structural model with bond price, Arora, Bohn, and Zhu (2005) compared model performance in predicting credit risk with CDS spreads. The CDS market gives new possibilities in estimating defaults and recoveries because the market price of CDS should reflect market assessment of credit events such as defaults.

This subsection will discuss some general concepts on CDS market and pricing of CDS spreads. A CDS is a contract ensuring protection against default. CDS contracts at several matu-

rity points between 1 and 10 years have been actively traded for several years. CDS with maturity T_{cds} is a contract involving protection between buyer and seller. If the reference company defaults at $\tau < T_{cds}$, the protection seller pays the buyer as much as a notional amount minus a recovery rate. This protects the protection buyer from the default event of the reference company. In turn, the protection buyer must make a periodic (usually every quarter) payment to the seller. This amount of periodic payment is called the CDS spread or CDS premium.

The fair annualized CDS spread with maturity T_{cds}, referred to as CDS_{Model}, is determined when the present value of the protection buyer's payment and the present value of the protection seller's payment are equal. In discrete-time frameworks, we assume that default occurs only in the middle of the default premium payment dates t_{j-1} and t_j. First, the present value of the protection buyer payment, $PV(Fixed\ Payment)$ is

$$PV(Fixed\ Payment) = CDS_{Model} d \sum_{j=1}^{M} DFD(0,t_j) q(0,t_j) + CDS_{Model} d/2 \sum_{j=1}^{M} DFD(0,t_j)(q(0,t_{j-1}) - q(0,t_j)), \tag{3.4}$$

where $t_j - t_{j-1} = T/M$; $DFD(0,t_j)$ is the default-free discount rate from time t_j to 0; $q(0,t_j)$ is the survival probability up to time t_j; and d is the premium paid fraction (quarterly=0.25). Equation (3.4) is comprised of two parts: the first part is the discounted premium if no default occurs and the second part is the accrued premium payment if default occurs between payment dates.

Second, the present value of the protection seller's payment $PV(Contingent\ Payment)$ is

$$PV(Contingent\ Payment) = (1-R) \sum_{j=1}^{M} DFD(0,t_j)(q(0,t_{j-1}) - q(0,t_j)), \tag{3.5}$$

where R is the recovery rate. The recovery rate means the fraction of the defaulted notional is recovered, so when default occurs, the protection seller pays $(1-R)\%$ of the protection buyer's loss.

CDS_{Model} is now determined where $PV(Contingent\ Payment) = PV(Fixed\ Payment)$ as follows:

$$CDS_{Model} = \frac{(1-R)\sum_{j=1}^{M} DFD(0,t_j)(q(0,t_{j-1}) - q(0,t_j))}{d\sum_{j=1}^{M} DFD(0,t_j) q(0,t_j) + d/2 \sum_{j=1}^{M} DFD(0,t_j)(q(0,t_{j-1}) - q(0,t_j))}. \tag{3.6}$$

Given a constant interest rate r [2], to calculate CDS_{Model}, we need a survival probability function $q(0,t)$. Survival probability calculations depend on the model assumptions, which will be given in section 3.5 and 3.6, respectively, under our two different model settings. More details about the CDS market and pricing of CDS are presented by Duffie and Singleton (1999), Pan and Singleton (2008), Hull and White (2000) and Lando (2004).

3.3 Preliminary Study: Variables and Model Structure Selection

As a study of correlated defaults, we apply the common factor method to two credit risk models: the structural and the intensity credit risk model. The concept behind a common factor is incorporating a correlation structure by assuming that the variables determining the default of firms share time-varying common factors. In order to select these common factors and to decide how we incorporate them into the model, we begin by conducting some preliminary studies.

In this section, as a proxy of a firm's defaultable bond, we examine stock price information for variables and model structure selection. Because stock prices are easily accessable information relative to other market prices representing a firm's financial health or its default risk, the preliminary results are based on a regression analysis of equity value. Furthermore, the stock price is a continuous variable, so it can be examined in a simple way rather than by rarely observed firm defaults. The equity values are measured by the product of stock prices and the number of outstanding.

3.3.1 Variables (Observed Factors) Selection

In order to choose financial variables that describe economic conditions, we refer to studies on stock return predictability. There are numerous papers showing that stock return can be predicted by financial variables, and various estimation methods and the prediction power of financial variables are covered (Campbell and Yogo (2006); Fama and French (1996a); Fama and French (1996b)). Even though there are differences in name and the form of the variables, market return,

[2] A simple arbitrage argument (Duffie and Singleton (2003)) shows the prices of these bonds are not highly sensitive to the level of interest rates.

value premium, and size premium are most commonly used factors, and they are relatively easily accessible by the public.

In this essay, as observed common covariates describing general economic condition, the following three variables are tested. As in previous studies, we do not include any interaction effects among variables.[3]

- Market Return (SP500) - log return of S&P500 ($\ln \frac{S\&P500_t}{S\&P500_{t-1}}$)
- Value Premium (HML) - The difference between the return on a portfolio of high-book-to-market stocks and the return on a portfolio of low-book-to-market stocks. This factor is applied after log-transformation as $\ln(1+HML)$.
- Size Premium (SMB) - The difference between the return on a portfolio of small stocks and the return on a portfolio of large stocks. This factor is applied after log-transformation as $\ln(1+SMB)$.

All these three variables are easily observable in the financial market. Market return represents the general economic condition, and HML is known as a proxy for relative distress. SMB represents the small stocks, which are not captured by the average market return.

For 109 companies in the data set, after treating log return of daily equity $\ln(equity_t) - \ln(equity_{t-1})$ as the response variable, each firm's regression model is fitted with three common factors. For all i, we fit the following regression models:

$$\ln S_{it} - \ln S_{i,t-1} = a_{0i} + a_{1i} SP500_t + a_{2i} HML_t + a_{3i} SMB_t + \varepsilon_{it}, \tag{3.7}$$

where i refers to firm and t refers to time, so $i = 1,...N(=109)$ and $t = 1,...T$. S_{it} is the equity price of firm i at time t, and ε_{it} is assumed to be independently and normally distributed. Details on 109 companies in our data set will be discussed in section 3.9.

Table 3.1 shows summary statistics of p-values for all 109 a_{0i}, a_{1i}, a_{2i} and a_{3i}. The most significant common factor is market return (SP500). Excepting one large extreme (p-value of

[3] The simple (completely pooled) regression results also show that interaction terms are all insignificant (p-values are 0.26,0.51,0.54, respectively).

Table 3.1: **Summary Statistics of Each Coefficients' p-values in the Preliminary Analysis.** For all 109 companies, the regression analysis of equity values with observed covariates are conducted independently as described in equation (3.7). These are summary statistics of p-values of a_{0i}, a_{1i}, a_{2i} and a_{3i} for all $i = 1, ...109$.

Covariates	Intercept	SP500	HML	SMB
Min	0.07023	3.284e-46	2.741e-20	1.188e-06
1st Qu.	0.37250	1.031e-23	1.815e-03	7.981e-03
Median	0.55362	4.674e-18	4.580e-02	2.515e-01
3rd Qu.	0.79997	2.290e-13	3.204e-01	5.244e-01
Max	0.99257	6.612e-01	9.957e-01	9.952e-01

AIG company is 0.6), SP500 is very significant for all other firms, while size premium (SMB) is not significant for most of the firms. Therefore, SP500 and value premium (HML) are adopted as our observed common factors in later analysis, which will be described from section 3.4.

3.3.2 Random Coefficients

Figure 3.1 displays the coefficients of market return (a_{1i}) and their 95% confidence intervals for all firms $i = 1, ...109$. This result shows that it is difficult to assume that market return effects are all the same across companies as Duffie, Eckner, Horel, and Saita (2009) assumed (the F-statistics which tests a completely pooled model versus completely independent models, is 15.1 and the p-value is < 0.0001). However, some of the firms still gathered near the completely pooled market return effect, which is 1.077.

The random coefficient idea compromises between two extreme assumptions: One extreme assumption is that general economic conditions affect all firms equally and the other assumption is that general economic conditions affect all firms differently. By adopting random coefficients and assigning a probability distributions on them, the results of individual companies can be partially pooled. The parameters governing assigned probability distributions on random coefficients (hyperparameters) are also estimated from the data (Gelman and Hill (2007)); thus we can measure the degree of connectivity between the coefficients of different firms for each of the common factors. Moreover, comparison among several factor effects under each correlated default model (structural or intensity) and comparison of certain factor effects between two different models are

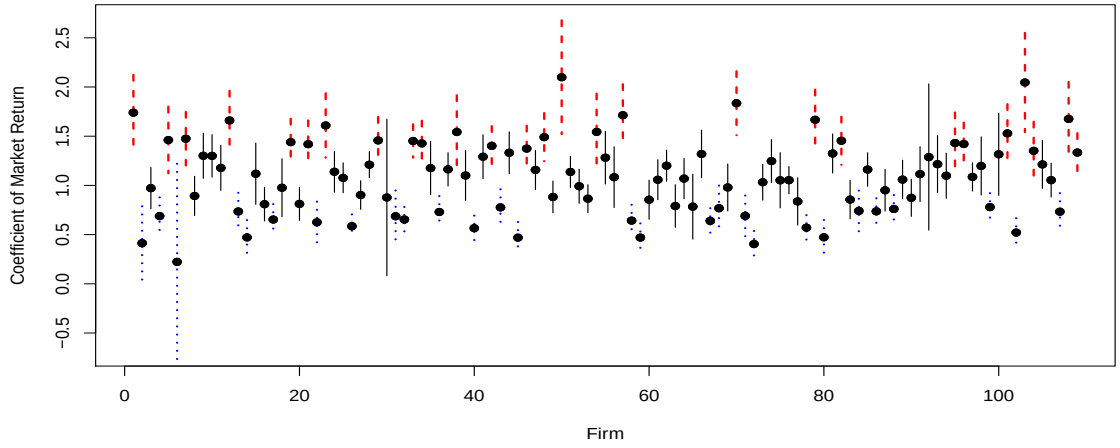

Figure 3.1: **Estimated Market Return Effects and Their 95% Confidence Interval (CI) in Preliminary Analysis.** This figure shows all estimated a_{1i} in equation (3.7) and their 95% CI. Dotted blue lines refer to 95% CI of companies with relatively small market return effect (lower 25%). Dashed red lines refer to 95% CI of companies with relatively large market return effect (upper 25%) The rest of the companies' 95% CIs are marked with solid black lines.

all possible through hyperparameter comparisons.

3.3.3 Random Time Effect (Hidden Factor)

In the previous subsections, the usefulness of fixed effects and necessity of random coefficients are explored. In this subsection, we will discuss whether they are enough to capture default correlation. Duffie, Eckner, Horel, and Saita (2009) addressed the insufficiency of observed common effect in capturing portfolio losses, especially high-quantile portfolio. Our preliminary study also shows the necessity of the hidden factor in correctly capturing default correlation.

Based on the regression analysis described in equation (3.7), the estimated correlation coefficient of log-equity return between two different firms i and j is defined as below:

$$\rho_{i,j}^{(regression)} = \frac{\sum_{m=1}^{3} \hat{a}_{mi}\hat{a}_{mj} Var(X_m) + \sum_{m=1}^{3}\sum_{m' \neq m} \hat{a}_{mi}\hat{a}_{m'j} Cov(X_m, X_{m'})}{\sigma_i \sigma_j}, \tag{3.8}$$

where i and j are indices for the firms and m is the index for the covariates; σ_i is the standard deviation of the i^{th} firm's log-equity return, and X_1, X_2, X_3 are the common factors SP500, HML,

SMB, respectively.

To determine how well the three observed covariates work in terms of capturing correlation between equity returns, we compare correlation coefficients derived from the regression model $\rho_{i,j}^{(regression)}$ with directly-estimated correlation coefficients. Directly-estimated correlation coefficients of log-equity returns $\rho_{i,j}^{(direct)}$ can be calculated using time-series of log-equity returns.

Figure 3.2 is a plot of $\rho_{i,j}^{(regression)}$ and $\rho_{i,j}^{(direct)}$. Most of the correlation coefficients are well fit by the regression model, but high correlation coefficients are generally under-estimated by the given observed common factors. Pairs with higher correlations are more downward biased as described in Duffie, Eckner, Horel, and Saita (2009). In other words, in equation (3.7), positive correlations remain between ε_{it} for these pairs. Because of this uncaptured correlation structure, the hidden factor is added into our model as a random time effect.

The next section summarizes the model structure that has been built based on these preliminary results. SP500, HML, and hidden factors will be used as common factors, and the form of their coefficients will be random.

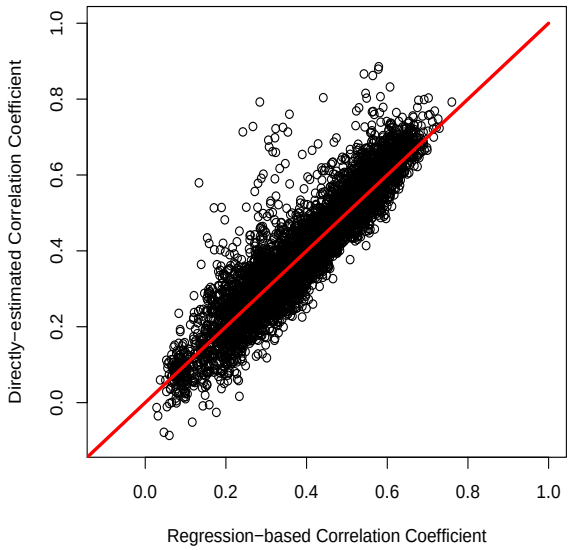

Figure 3.2: **Scatter Plot of $\rho_{i,j}^{(regression)}$ and $\rho_{i,j}^{(direct)}$.** $\rho_{i,j}^{(regression)}$ are derived from regression results with observed covariates, SP500, HML, SMB, as shown in equation (3.8). $\rho_{i,j}^{(direct)}$ are calculated directly using time-series of log-equity returns. The solid red line is $y = x$ line.

3.4 Overview of the Models

Our approach is different from previous works: we build two different credit risk models with a parallel frame by combining ideas in previous structural and intensity models research. In order to facilitate better understanding, this section provides a general overview of the parallel frame based on which two correlated credit risk models are derived: (1) a correlated structural credit risk model with common factors for the asset value process: and (2) a correlated intensity credit risk model with common factors for the default intensity process.

The parallel frame of the models in this essay can be summarized with the following three points: First, a state-space model is used here. Because asset and default intensity processes are not observable, they are assumed to be state variables that determine a corresponding firm's default time and other financial market prices, such as stocks, bonds, CDS prices and so on. Second, the state variables are assumed to follow a stochastic process. By doing so, empirical analysis does not rely on large amount of data to ensure that the data contains a large enough number of defaults, as previous research has suggested. Third, by assuming that state processes share common factors with random coefficients, a default correlation structure is incorporated. Based on preliminary results, we model state variables with hierachical linear models with random coefficients and random time effects (hidden factor). The hidden factor is expected to adjust bias in default distribution estimation under previous without-hidden-factor models. Random coefficients make it possible to assume non-equal factor effects across the firms.

In mathematical form, the common frame of both our correlated structural and intensity models is as below:

$$(State\ equation) \quad : \quad d\Omega_{it} = \kappa_\mu(\Omega_{it},\Theta)dt + \kappa_\sigma(\Omega_{it},\Theta)dW_{it}^{\Omega}, \tag{3.9}$$

$$dW_{it}^{\Omega} = a_i dX_t + b_i dY_t + \sqrt{1-a_i^2-b_i^2}\,dW_{it}^{\varepsilon}, \tag{3.10}$$

$$(Observation\ equation) \quad : \quad \omega_{it} = g_\omega(\Omega_{it},t,\Theta) + Z_{it}^{\omega}, \tag{3.11}$$

where $|a_i| < 1$ and $0 < b_i < 1$. Ω_{it} refers to state variables, which determine firm i's default at time t. For the structural model, Ω_{it} are the firm's asset values, V_{it}, and for the intensity model,

Ω_{it} are the firm's default intensities, λ_{it}. $\kappa_\mu()$ and $\kappa_\sigma()$ are the functions for determining the mean and volatility levels of the Ω_{it} process, and Θ is a parameter vector determining mean and volatility of the stochastic process of state variables. X_t is the observed common factor vector (SP500 and HML), and Y_t is the hidden common factor; these are all assumed to be independent. In applying the observed common factor, Brownian motion, which underlies the evolution of the corresponding observed common factors, is used. All Ws are standard Brownian motions; W_{it}^{Ω}s are not independent but W_{it}^{ε}s are assumed to be independent for all firms i.

Here, ω_{it} is the financial derivative price, which is observed in the market. The function $g()$ is a function of theoretical prices (derived from a certain pricing model) of ω. Z_{it}^{ω} refers to market noise of ω. As observations ω_{it}, we adopt equity price and CDS spreads in this essay.

3.5 Model 1: Correlated Structural Model

This section introduces model details under structural credit risk model settings. The two models in this essay are built under a parallel frame described in the previous section; model details are enumerated by different levels of equations.

3.5.1 State Level: Correlated Asset Model

Under the structural model, in order to capture asset value correlation, the hidden factor idea is newly applied to structural model settings. N different companies are considered, and V_{it} refers to the i^{th} company's asset value at time t. V_{it} is assumed to follow geometric Brownian motion under objective-measure P as below:

$$dV_{it} = \mu_i V_{it} dt + \sigma_i V_{it} dW_{it}^{V}, \tag{3.12}$$

where $\sigma_i > 0$.

The dependence structure between firms is modeled by correlated standard Brownian motion

W_{it}^V as follows:

$$dW_{it}^V = a_{1i}dX_{1t} + a_{2i}dX_{2t} + b_i dY_t + \sqrt{1 - a_{1i}^2 - a_{2i}^2 - b_i^2}dW_{it}^{\varepsilon}, \tag{3.13}$$

where $|a_{1i}| < 1$,$|a_{2i}| < 1$ and $0 < b_i < 1$. The absolute values of a_{1i}, a_{2i} and b_i values fall between 0 and 1 because of the square-root term in equation (3.13), and b_i is assumed to be positive for the separate identification of b_i and Y_t. Common factor processes, dX_{1t} and dX_{2t}, are Brownian motion underlying the SP500 and HML variables, respectively.[4]

In applying the observed common factor, we use Brownian motion underlying the evolution of corresponding observed common factors. The hidden common factor Y_t is also assumed to follow Brownian motion because this can be interpreted in the view of a missing common factor. Volatility of X_{1t}, X_{2t}, and Y_t are all set equal to 1 because they can be measured with each coefficient term a_{1i}, a_{2i} and b_i, respectively. All common factors are assumed to be independent.

As mentioned in introduction, one of the extensions from previous studies in this essay is the random coefficient assumption. All coefficients a_{1i}, a_{2i} and b_i and μ_i for $i = 1, ...N$ are connected, respectively, under the adoption of random coefficients. The distribution assumptions for each parameters are:

- $\mu_i \sim Normal(\mu_\mu, \sigma_\mu)\ for\ i = 1...N$,
- $(a_{1i}, a_{2i}, b_i) \sim P(a_{1i}, a_{2i}, b_i)$ for $i = 1...N$: Product of three truncated normals with constraints

$$\begin{aligned} P(a_{1i}, a_{2i}, b_i) \quad \propto \quad & \exp\{-\frac{(a_{1i} - \mu_{a_1})^2}{2\sigma_{a_1}^2}\} I_{-1 \le a_{1i} \le 1} \\ & \times \exp\{-\frac{(a_{2i} - \mu_{a_2})^2}{2\sigma_{a_2}^2}\} I_{-1 \le a_{2i} \le 1} \\ & \times \exp\{-\frac{(b_i - \mu_b)^2}{2\sigma_b^2}\} I_{0 \le b_i \le 1} \\ & \times I_{a_{1i}^2 + a_{2i}^2 + b_i^2 \le 1} \end{aligned}$$

Because all a_{1i}, a_{2i} and b_i are bounded and jointly constrained due to the square-root term in

[4] Instead of using return values, we transform all observed common factor to price value.

equation (3.13), truncated normal distributions with joint constraint for a_{1i}, a_{2i} and b_i are adopted as a joint prior distribution.

For the prior distribution of hyperparameters μ_μ and σ_μ, we use the non-informative prior as follows,

$$(\mu_\mu, \mu_\sigma) \propto \frac{1}{\sigma_\mu}.$$

For $\mu_{a_1}, \sigma_{a_1}, \mu_{a_2}, \sigma_{a2}, \mu_b$ and σ_b, we use a flat (uniform) hyperprior within the reasonably bounded interval. We use informative prior for $\mu_{a_1}, \sigma_{a_1}, \mu_{a_2}, \sigma_{a2}, \mu_b$ and σ_b, because of the properness of posterior distribution. For the prior of hyper-mean parameters, we use uniform distribution for $(-1000, 1000)$ interval. For the prior of hyper-standard deviation parameters, we use uniform distribution for $(0, 100)$. Because, a_{1i}, a_{2_i} and b_i are constraint in the (-1,1) or (0,1) interval, mean and standard deviation beyond these intervals which hyperpriors are defined on, are not reasonable.[5]

3.5.2 Observation Level: Pricing Function of Equity and CDS Prices

Another distinctive approach in our study is that both equity and CDS prices are used together in the observation equation. As mentioned in section 3.3, equity price is derived from stock market information as a product of stock price and number of outstanding. We define S_{it} as observed equity price of the i^{th} firm at time t, and CDS_{it} as observed CDS spread of the i^{th} firm at time t. With stock market noise $Z_{S_{it}}$ and CDS market noise $Z_{CDS_{it}}$, observation equations are:

$$\ln S_{it} = g_S(V_{it}, \sigma_i, T_D, D) + Z_{it}^S, \tag{3.14}$$

$$\ln CDS_{it} = g_{CDS}(V_{it}, \sigma_i, T_{CDS}) + Z_{it}^{CDS}, \tag{3.15}$$

[5] We tried non-informative hyperpriors and beta distributions for priors, but there is no big differences in estimation results.

where

$$Z_{it}^{S} \sim N(0, \sigma_S), \tag{3.16}$$

$$Z_{it}^{CDS} \sim N(0, \sigma_{CDS}). \tag{3.17}$$

T_D and T_{CDS} are the maturity of bond and CDS and D is the face value of bond. $g_S()$ and $g_{CDS}()$ represent model-derived (theoretical) equity and CDS values, and will be explained in more detail in the following two subsections.

For an error between model and market values, we assume a normal distribution for both stock and CDS, as in Duan and Fulop (2009) and Pan and Singleton (2008), respectively. σ_S and σ_{CDS} are assumed to be known because with a single process of equity or 5-year CDS prices, it is hard to separate error effect and unknown asset value effects within them. The standard deviation of stock market noise σ_S is set to be 0.01. The assumption that the 90% confidence interval for the difference between model derived value and market value is around $\pm 1.645\%$ seems too strict, but according to empirical results in Duan and Fulop (2009), this level is plausible.[6]

For the CDS market, a larger variance value is set than for the stock market due to different market liquidity. To compare market liquidity, bid-ask spreads are usually used because it is well known that there is positive relationship between bid-ask spread and market noise (Ait-Sahalia and Yu (2009)). During 2009, the average CDS bid-ask spread of 125 companies comprising CDX.NA.IG13 was 10.4 basis points. The average stock prices bid-ask spread of the same 125 companies was 0.87. Hence, just as Pan and Singleton (2008) assumed a different standard deviation of CDS markets noise as a value proportional to their market bid-ask spreads, σ_{CDS} is set here to be 0.1, which is a value 10 times larger than σ_S.

Even though additional assumptions on the noise terms are necessary, a scheme to adopt market noise terms Z_{it}^{S} and Z_{it}^{CDS} into a model has meaning in two ways. First, in the sense of the statistical model, market noise makes it possible to combine the information from several differ-

[6] According to Duan and Fulop (2009), the mean of standard deviations of 100 randomly chosen non-DowJones30 companies' stock market noise is 0.0043 and the 90 percentile is 0.016 under Merton's structural model assumption. The mean of standard deviations of DowJones30 companies' stock market noise is 0.003 and the 90 percentile is 0.007.

ent markets in estimating the asset values process. Second, in the sense of the financial theory, market noises capture the model mis-specification error or short-term discrepancies in the supply and demand within financial markets. More on market noise and model mis-specification errors in stock market are discussed in the Chapter 1.

Equity Valuation under the Black-Cox Structural Model: Functional Form of $g_S()$

To determine a functional form of $g_S()$ which is model-derived (theoretical), the Black-Cox model (Black and Cox (1976)) is used. The Black-Cox model has more relaxed assumptions on default time than the Merton model. However, because model mis-specification error is allowed, it might not be important which equity pricing model is used to derive the theoretical equity price. Black and Cox consider a process for asset value, which under the risk-neutral measure is

$$dV_{it} = rV_{it}dt + \sigma_i V_{it} dW_{it}^{QV}, \tag{3.18}$$

where r is a known constant risk-less rate. When asset value falls below a certain boundary L_t, the firm defaults. The standard default barrier function in the Black-Cox model is adopted as below:

$$L_{it} = L_{i0} \exp\left(-\gamma(T_D - t)\right), \tag{3.19}$$

for some positive constants L_{i0} and γ. We assume same γ across all the firms. The basics in pricing debt and equity under the Black-Cox model settings are introduced in the Chapter 1.2.

The functional form of $g_S()$ in equation (3.14) is

$$g_S(V_{it}, \sigma_i, T_D, D) = \ln\left(V_{it} - B_{BC}(V_{it}, t, T_D, D)\right), \tag{3.20}$$

where $B_{BC}(V_{it}, t, T_D, D)$ is the price of the defaultable bond that pays D if the bond default time $\tau > T_D$. Details on $B_{BC}(V_{it}, t, T_D, D)$ is in the appendix A.1. The maturity of bond T_D and the face value of the bond D are all known constants.

Survival Probability under the Structural Model: Functional Form of $g_{CDS}()$

The functional form of $g_{CDS}()$ in equation (3.15) is determined by equation (3.6) with the following survival probability function. Based on Black and Cox (1976), Zhou (2001) gives a closed-form solution for default probability under the structural credit risk model.

When the dynamic of asset follows geometric Brownian motion as shown in equation (3.18), the survival probability $q_i(0,t)$ of firm i by time t is

$$q_i(0,t) = 1 - \Phi(-Z/\sqrt{t} - \frac{\phi_i - \gamma}{\sigma_i}\sqrt{t}) - \exp(\frac{2(\gamma-\phi_i)Z_i}{\sigma_i})\Phi(-Z_i/\sqrt{t} + \frac{\phi_i - \gamma}{\sigma_i}\sqrt{t}), \qquad (3.21)$$

where $\phi_i = r - (\sigma_i)^2/2$, $Z_i = \frac{\ln(V_{i0}/K_i)}{\sigma_i}$, $K_i = L_{i0}(\exp(\gamma T_D))$ and Φ is the cumulative distribution function (cdf) of the standard normal distribution. L_{i0} and γ are default barrier parameters described in equation (3.19).

Therefore, the functional form of $g_{CDS}()$ is

$$g_{CDS}(V_{it}, \sigma_i, T_{CDS}) = \ln\left(CDS_{StrucModel}(V_{it}, \sigma_i, T_{CDS})\right), \qquad (3.22)$$

where

$$CDS_{StrucModel}(V_{it}, \sigma_i, T_{CDS}) = \frac{(1-R_i)\sum_{j=1}^{M} DFD(0,t_j)(q_i(0,t_{j-1}) - q_i(0,t_j))}{d\sum_{j=1}^{M} DFD(0,t_j)q_i(0,t_j) + d/2\sum_{j=1}^{M} DFD(0,t_j)(q_i(0,t_{j-1}) - q_i(0,t_j))}$$

where $q_i(0,t)$ is determined by the value of V_{it} and σ_i, as given in equation (3.21). The maturity of CDS T_{CDS} is a known constant.

3.5.3 Properties of the Model

In summary, we build the following state-space model for the correlated structural model.

$$
\begin{aligned}
(State\ equation) \quad &: \quad d\ln V_{it} = (\mu_i - \frac{\sigma_i^2}{2})dt + \sigma_i(a_{1i}dX_{1t} + a_{2i}dX_{2t} + b_i dY_t) \\
&\qquad\qquad +\sigma_i\sqrt{1-a_{1i}^2-a_{2i}^2-b_i^2}dW_{it}^{\varepsilon} \\
(Observation\ equation) \quad &: \quad \begin{cases} \ln S_{it} = g_S(V_{it},\sigma_i,T_D,D) + Z_{it}^S \\ \ln CDS_{it} = g_{CDS}(V_{it},\sigma_i,T_{CDS}) + Z_{it}^{CDS} \end{cases}
\end{aligned}
$$

The state equation is a hierachical mixed effects linear model with random coefficients and random time effects. At the observation level, the Black-Cox model is assumed, but by allowing noise term, we incorporate the model with mis-specification error.

In a discrete-time framework, the conditional distribution $P_V(\ln V_{it}|\ln V_{it-1}, X_{1:t}, Y_{1:t}, \Theta_V)$ for all i and t are assumed independently and identically distributed Normal, with the following mean $\mu_{structural}$ and standard deviation $\sigma_{structural}$:

$$
\begin{aligned}
\mu_{structural} &= \ln V_{it-1} + (\mu_i - \frac{\sigma_i^2}{2})\Delta t + \sigma_i(a_{1i}(\Delta X_{1t}) + a_{2i}(\Delta X_{2t}) + b_i(\Delta W_t)), \qquad (3.23) \\
\sigma_{structural} &= \sigma_i\sqrt{1-a_{1i}^2-a_{2i}^2-b_i^2}\sqrt{\Delta t}, \qquad (3.24)
\end{aligned}
$$

where parameter vector Θ_V includes all random coefficients $(\mu_1,...\mu_N, a_{11},...,a_{1N}, a_{21},...,a_{2N}, b_1,...b_N)$ and their hyperparameters $(\mu_\mu, \sigma_\mu, \mu_{a_1}, \sigma_{a_1}, \mu_{a_2}, \sigma_{a_2}, \mu_b, \sigma_b)$.

One of the advantages of our model is that asset process volatility can be decomposed into common and idiosyncratic parts:

- Common part: $\sigma_i\sqrt{a_{1i}^2+a_{2i}^2+b_i^2}$
- Idiosyncratic part: $\sigma_i\sqrt{1-a_{1i}^2-a_{2i}^2-b_i^2}$

Furthermore, through our model, the correlation between the i^{th} firm and j^{th} firm can be mea-

sured with coefficients a_1, a_2, and b:

$$Corr(d\ln V_{it}, d\ln V_{jt}|a_1, a_2, b, \sigma_i, \sigma_j) = a_{1i}a_{1j} + a_{2i}a_{2j} + b_i b_j. \tag{3.25}$$

3.6 Model 2 : Correlated Intensity Model

This section introduces, our model details under intensity credit risk model settings. A similar format is employed for model 2 as for model 1.

3.6.1 State Level : Correlated Intensity Model

Duffie, Eckner, Horel, and Saita (2009) directly estimated default intensities from an extensive history of defaults using the factor method. However, in applying this, a large data set, both in terms of time points and number of firms, is necessary because the default event itself is very rare. This leads us to specify default intensity as a latent variable and observable through a pricing function as was done for the asset process in the structural model.

We assume that i^{th} firm default intensity is expressed with CIR type specifications[7]. Simple normal or Vesicek type process assumptions are mathematically easy, but it do not guarantee the non-negativeness of intensities. Further, the lognormal distribution does not have a closed-form solution for survival probability, which is essential in CDS pricing.

Under CIR specifications, the dynamics for the default intensity processes, λ_{it}, under risk neutral measure Q, are given by

$$d\lambda_{it} = (\mu_i' - \lambda_{it})dt + \sigma_i'\sqrt{\lambda_{it}}dW_{it}^{\lambda}. \tag{3.26}$$

These intensity dynamics already ensure that the intensities are always nonnegative.

Similar to Model 1, the dependence structure between firms are modeled with correlated W_{it}^{λ}

[7] Longstaff and Rajan (2008), Bielecki, Jeanblanc, and Rutkowski (2011), Duffee (1999), and Brigo and Alfonsi (2005) assumed a CIR type intensity process behavior. Pan and Singleton (2008) assumed a lognormal distribution. Liang, Ma, Wang, and Ji (2011) assumed a Vasicek type intensity behavior.

as follows:

$$dW_{it}^{\lambda} = a'_{1i}dX_{1t} + a'_{2i}dX_{2t} + b'_i dY_t + \sqrt{1 - a'^2_{1i} - a'^2_{2i} - b'^2_i}dW_{it}^{\varepsilon}. \tag{3.27}$$

Observed common factors X_1 and X_2, hidden common factor Y, and W_{it}^{ε} are defined the same as in Model 1. To differentiate intensity model from structural model, we use prime notation for all parameters $\mu'_i, \sigma'_i, a'_{1i}, a'_{2i}$ and b'_i.

Related to the mean reverting nature of the intensity dynamic, we assume that all mean reverting rates are equal to 1 because, with only one year data on which we are focusing, it is difficult to specify mean reverting rate and volatility separately. With fixed mean reverting rate, excessively or insufficiently captured mean reverting nature is adjusted by σ'_i estimates.[8]

We applied same prior conditions as in Model 1.

3.6.2 Observation Level : Pricing Function of CDS Spread

With CDS market noise $Z_{it}^{CDS\prime}$, observation equations are as follows:

$$\ln(CDS_{it}) = l_{CDS}(\lambda_{it}, \mu'_i, \sigma'_i, T_{CDS}) + Z_{it}^{CDS\prime}, \tag{3.28}$$

$$Z_{it}^{CDS\prime} \sim Normal(0, \sigma_{CDS}), \tag{3.29}$$

where $\sigma_{CDS} = 0.01$ and the functional form of $l()$, which is the theoretical CDS price under the intensity model settings, is shown in the following subsection.

We assume σ_{CDS} is known, as in our correlated structural model. However, different from the structural model, we cannot adopt equity prices under the intensity model setting, so it is impossible to define σ_{CDS} as proportional to the relatively well-known σ_S. Hence, in line with the concept of a parallel frame between two models, we assume that CDS prices provide information on the intensity as much as or at least as equity prices do on the asset; this occurs because we can

[8] The stationary distribution of the CIR process is gamma distribution with shape parameter $\frac{2k\mu_{\lambda_i}}{\sigma'^2_i}$ and scale parameter $\frac{\sigma'^2_i}{2k}$, where k is the mean reverting rate. It is well known that the mean reverting rate of default intensities is small (Brigo and Alfonsi (2005), Liang, Ma, Wang, and Ji (2011)). Longstaff and Rajan (2008) simplified their model by assuming in all cases $\mu_{\lambda_i} = k_{\lambda_i} = 0$, and they showed that the general case of the model offers only marginal improvement.

only rely on the CDS prices under the intensity model setting.[9]

Survival Probability under the Intensity Model: Functional Form of $l()$

The functional form of $l()$ is determined by equation (3.6) with the following survival probability function. By Hull and White (2000) and Lando (2004), the survival probability $q(0,t)$ up to time t in a CIR-setting has a closed-form solution.

When the intensity process follows CIR processes as in equation (3.26), we obtain the survival probability as follows:

$$\begin{aligned} q_i(0,t) &= E\exp(-\int_0^t \lambda_{is}ds) \\ &= \exp(A(i,t)+B(i,t)\lambda_0), \end{aligned} \tag{3.30}$$

where

$$\begin{aligned} \delta_i &= \sqrt{1+2(\sigma_i')^2}, \\ A(i,t) &= \frac{2\mu_i'}{(\sigma_i')^2}\ln\frac{2\delta_i\exp((\delta_i+1)t/2)}{2\delta_i+(\delta_i+1)(\exp(\delta_i t)-1)}, \\ B(i,t) &= \frac{-2(\exp(\delta_i t)-1)}{2\delta_i+(\delta_i+1)(\exp(\delta_i t)-1)}. \end{aligned}$$

Therefore, the functional form of $l()$ is

$$l(\lambda_t,\mu_i',\sigma_i',T_{CDS}) = \ln\left(CDS_{IntenModel}(\lambda_{it},\mu_i',\sigma_i',T_{CDS})\right), \tag{3.31}$$

where

$$CDS_{Inten-Model}(\lambda_{it},\mu_i',\sigma_i',T_{CDS}) = \frac{(1-R_i)\sum_{j=1}^{M} DFD(0,t_j)(q_i(0,t_{j-1})-q_i(0,t_j))}{d\sum_{j=1}^{M} DFD(0,t_j)q_i(0,t_j)+d/2\sum_{j=1}^{M} DFD(0,t_j)(q_i(0,t_{j-1})-q_i(0,t_j))}$$

and $q_i(0,t)$ is determined by the value of $\lambda_{it},\mu_i',\sigma_i'$ as given in equation (3.30).

[9] In usual previous research, instead of adopting model mis-specification or market error, the 5-year CDS contract was usually assumed to be priced perfectly, so that the pricing function could be inverted for intensity in the customary intensity model and CDS price studies (Pan and Singleton (2008))

3.6.3 Properties of the Model

Because the two models have similar settings, the properties of intensity model can be summarized in a similar manner to Model 1.

First, we build the following state-space model.

$$\begin{aligned}
(State\ equation) \quad &: \quad d\lambda_{it} = (\mu_i' - \lambda_{it})dt + \sigma_i'\sqrt{\lambda_{it}}(a_{1i}' dX_{1t} + a_{2i}' dX_{2t} + b_i' dY_t) \\
&\qquad\qquad + \sigma_i'\sqrt{\lambda_{it}}\sqrt{1 - a_{1i}'^2 - a_{2i}'^2 - b_i'^2}\, dW_{it}^{\varepsilon} \\
(Observation\ equation) \quad &: \quad \ln(CDS_{it}) = l(\lambda_{it}, \mu_i', \sigma_i', T_{CDS}) + Z_{it}^{CDS\prime}
\end{aligned}$$

Second, in the discrete-time framework, the conditional distribution $[\lambda_{it}|\lambda_{it-1}, X_{1:t}, Y_{1:t}, \Theta_\lambda]$ is identically and independently distributed Normal with the following $\mu_{intensity}$ and $\sigma_{intensity}$:

$$\mu_{intensity} = \lambda_{it-1} + (\mu_i' - \lambda_{it})\Delta t + \sigma_i'\sqrt{\lambda_{it-1}}(a_{1i}'(\Delta X_{1t}) + a_{2i}'(\Delta X_{2t}) + b_i'(\Delta Y_t)), \quad (3.32)$$

$$\sigma_{intensity} = \sigma_i'\sqrt{\lambda_{it-1}}\sqrt{1 - a_{1i}'^2 - a_{2i}'^2 - b_i'^2}\sqrt{\Delta t}. \quad (3.33)$$

Third, the volatility $\sigma_i'\sqrt{\lambda_{it-1}}$ of the i^{th} firm default intensity at time t is divided into two parts:

- Common part : $\sigma_i'\sqrt{\lambda_{it-1}}\sqrt{a_{1i}'^2 + a_{2i}'^2 + b_i'^2}$.
- Idiosyncratic part : $\sigma_i'\sqrt{\lambda_{it-1}}\sqrt{1 - a_{1i}'^2 - a_{2i}'^2 - b_i'^2}$.

Finally, the correlation between the i^{th} firm and the j^{th} firm is

$$Corr(d\lambda_{it}, d\lambda_{jt}|a_1', a_2', b', \sigma_i', \sigma_j') = a_{1i}' a_{1j}' + a_{2i} a_{2j}' + b_i' b_j'. \quad (3.34)$$

3.7 Estimation

Basic statistical tools used for parameter estimation are Gibbs sampling and the Metropolis Hasting (MH) algorithm. To improve convergence rates, the generalized Gibbs sampling technique is adopted for highly correlated parameters and unknowns. This section discusses our Bayesian estimation procedure under our two different models.

3.7.1 Estimation Procedure for Model 1: Correlated Structural Model

In our correlated structural model, the parameter vector we need to estimate is $\Theta_V = (\mu_1, ...\mu_N$ $, a_{11}, ...a_{1N}, a_{21}, ...a_{2N}, b_1, ...b_N, \sigma_1, ...\sigma_N, \mu_\mu, \sigma_\mu, \mu_{a_1}, \sigma_{a_1}, \mu_{a_2}, \sigma_{a_2}, \mu_b, \sigma_b)$, and $Y = Y_{1:T}$, $\ln V = \ln V_{1:N,1:T}$ are also unknown. Given that all $\ln V_{i0}$ are known, the conditional distribution of unknown values $P(\ln V, Y, \Theta_V | S, CDS, X)$ is

$$\begin{aligned} P(\ln V, Y, \Theta_V | S, CDS, X) &\propto P_{CDS,S}(CDS, S | \ln V, \Theta_V, X, Y) P_V(V | X, Y, \Theta_V) \times \\ &\quad P_Y(Y | \Theta_V) P_\Theta(\Theta_V) \\ &= P_S(S | \ln V, \Theta_V) P_{CDS}(CDS | \ln V, \Theta_V) P_V(V | X, Y, \Theta_V) \times \\ &\quad P_Y(Y | \Theta_V) P_\Theta(\Theta_V) \\ &= \prod_{i=1}^{N} \prod_{t=1}^{T} P_S(\ln S_{it} | \ln V_{it}, \Theta_V) P_{CDS}(CDS_{it} | \ln V_{it}, \Theta_V) \times \\ &\quad P_V(\ln V_{it} | \ln V_{it-1}, X, Y, \Theta_V) P_Y(Y_t | Y_{t-1} \Theta_V) P_\Theta(\Theta_V), \end{aligned} \tag{3.35}$$

where

$$P_V(\ln V_{it} | \ln V_{it-1}, X_{t-1:t}, Y_{t-1,t}, \Theta_V) \sim N(\mu_{structural}, \sigma_{structural}), \tag{3.36}$$

$$P_S(\ln S_{it} | \ln V_{it}, \Theta_V) \sim N(g_S(V_{it}, \sigma_{V_i}, T_D, D), \sigma_S), \tag{3.37}$$

$$P_{CDS}(\ln CDS_{it} | \ln V_{it}, \Theta_V) \sim N(g_{CDS}(V_{it}, \sigma_{V_i}, T_{CDS}), \sigma_{CDS}), \tag{3.38}$$

$$P_Y(\Delta Y_t | \Theta) \sim N(0, \Delta t). \tag{3.39}$$

$\mu_{structural}, \sigma_{structural}$ are described in equation (3.23) and (3.24). The $g_S()$ and $g_{CDS}()$ functions are the equity and CDS pricing solution and are given in equation (3.20) and equation (3.22) respectively. In order to make estimation simpler, instead of estimating Y_t process, ΔY_t is estimated given $Y_0 = 0$.

We divide stock and CDS market price distribution $P_S()$ and $P_{CDS}()$ separately in equation (3.35). It is difficult to assume that contemporary stock and CDS markets are independent. However, given the state variable that determines all financial markets, an independence assumption

on market noise seems plausible. Thus stock and CDS markets are assumed to be conditionally independent given asset values.

To obtain Monte Carlo samples from the joint posterior distribution, the conditional sampling steps are iterated with following target distributions:

- $\Delta Y_t \sim [\Delta Y_t | \ln V_{1:N,1:t}, \Theta_V, X_{1,1:t}, X_{2,1:t}, S_{1:N,1:t}, CDS_{1:N,1:t}]$
$$\propto \prod_{i=1}^{N} P_V(\ln V_{it} | \ln V_{it-1}, X_{1,t-1:t}, Y_{t-1:t}, \Theta_V) P_Y(\Delta Y_t | \Theta_V),$$

- $\ln V_{it} \sim [\ln V_{it} | \Theta_V, \Delta Y_t, X_{1,1:t}, X_{2,1:t}, S_{i,1:t}, CDS_{i,1:t}]$
$$\propto P_V(\ln V_{it} | \ln V_{it-1}, X_{1,t-1:t}, X_{2,t-1:t}, Y_{t-1,t}, \Theta_V) \times$$
$$P_V(\ln V_{i(t+1)} | \ln V_{it}, X_{1,t-1:t}, X_{2,t-1:t}, Y_{t-1,t}, \Theta_V) P_S(\ln S_{it} | \ln V_{it}, \Theta) P_{CDS}(\ln CDS_{it} | \ln V_{it}, \Theta)$$

- $\mu_i \sim [\mu_{V_i} | \ln V_{i,1:T}, \Theta_{V[-\mu]}, \Delta Y_{1:T}, X_{1,1:T}, X_{2,1:T}, S_{i,1:T}, CDS_{i,1:T}]$
$$\sim N\left(\frac{\frac{\mu_\mu}{\sigma_\mu^2} + \frac{T\overline{\ln V_{it}^*}}{\sigma_i^2(1-a_{1i}^2-a_{2i}^2-b_i^2)/\Delta t}}{\frac{1}{\sigma_\mu^2} + \frac{T}{\sigma_i^2(1-a_{1i}^2-a_{2i}^2-b_i^2)/\Delta t}}, \frac{1}{\frac{1}{\sigma_\mu^2} + \frac{T}{\sigma_i^2(1-a_{1i}^2-a_{2i}^2-b_i^2)/\Delta t}} \right)$$
where $\ln V_{it}^* = (lnV_{it} - \ln V_{i,t-1} - \sigma_i(a_{1i}\Delta X_{1t} + a_{2i}\Delta X_{2t} + b_i \Delta Y_t))/\Delta t + \sigma_i^2/2$, and
$\overline{\ln V_{it}^*} = \frac{\sum_{t=1}^{T} \ln V_{it}^*}{T}$

- $(a_{1i}, a_{2i}, b_i) \sim [a_{1i}, a_{2i}, b_i | \ln V_{i,1:T}, \Theta_{V[-(a_{1i},a_{2i},b_i)]}, \Delta Y_{1:T}, X_{1,1:T}, X_{2,1:T}, S_{i,1:T}, CDS_{i,1:T}]$
$$\propto \prod_{t=1}^{T} P_V(\ln V_{it} | \ln V_{it-1}, X_{1,t-1:t}, X_{1,t-1:t}, Y_{t-1:t}, \Theta_V) P_{a1}(a_{1i} | \mu_{a_1}, \sigma_{a_1}) \times$$
$$P_{a2}(a_{2i} | \mu_{a_2}, \sigma_{a_2}) P_b(b_i | \mu_b, \sigma_b) P_{ma1} P(\mu_{a1}, \sigma_{a1}) P_{ma2} P(\mu_{a2}, \sigma_{a2}) P_{mb} P(\mu_b, \sigma_b)$$

- $\sigma_i \sim [\sigma_{V_i} | \ln V_{i,1:T}, \Theta_{V[-\sigma]}, \Delta Y_{1:T}, X_{1,1:T}, X_{2,1:T}, S_{i,1:T}, CDS_{i,1:T}]$
$$\propto \prod_{t=1}^{T} P_V(\ln V_{it} | \ln V_{it-1}, X_{1,t-1:t}, X_{2,t-1:t}, Y_{t-1:t}, \Theta_V) P_S(\ln S_{it} | \ln V_{it}, \Theta_V) \times$$
$$P_{CDS}(\ln CDS_{it} | \ln V_{it}, \Theta_V) P_\sigma(\sigma_i)$$

The hyperparameter (μ_μ, σ_μ) is sampled with the following Gibbs sampling scheme:

$$\mu_\mu | \sigma_\mu, \mu_1, ...\mu_N \sim N(\bar{\mu}, \sigma_\mu/N), \tag{3.40}$$

$$\sigma_\mu | \mu_\mu, \mu_1, ...\mu_N \sim \textit{Inv-}\chi^2(N-1, S^2), \textit{ where, } S^2 = \frac{\sum_{i=1}^{N}(\mu_i - \bar{\mu})^2}{N-1}. \tag{3.41}$$

Hyperparameters (μ_{a_1},σ_{a_1}),(μ_{a_2},σ_{a_2}), and (μ_b,σ_b) are sampled with the MH scheme with the following target distributions:

$$\mu_{a_1},\sigma_{a_1} \sim [\mu_{a_1},\sigma_{a_1}|a_{11},...a_{1N}] \propto \prod_{i=1}^{N} P_{a1}(a_{1i}|\mu_{a_1},\sigma_{a_1})P_{ma1}(\mu_{a_1},\sigma_{a_1}), \tag{3.42}$$

$$\mu_{a_2},\sigma_{a_2} \sim [\mu_{a_2},\sigma_{a_2}|a_{21},...a_{2N}] \propto \prod_{i=1}^{N} P_{a2}(a_{2i}|\mu_{a_2},\sigma_{a_2})P_{ma2}(\mu_{a_2},\sigma_{a_2}), \tag{3.43}$$

$$\mu_b,\sigma_b \sim [\mu_b,\sigma_b|b_1,...b_N] \propto \prod_{i=1}^{N} P_b(b_i|\mu_b,\sigma_b)P_{mb}(\mu_b,\sigma_b). \tag{3.44}$$

As proposal distributions for other parameters and unknowns in implementing the MH algorithm, we use the following distributions(in j^{th} iteration).

- For Y_t: $p(Y_t^{(j)}|Y_t^{(j-1)}) \sim Normal(Y_t^{(j-1)},s0)$,
- For $\ln V_{it}$: $p(\ln V_{it}^{(j)}|\ln V_{it}^{(j-1)}) \sim Normal(\ln V_{it}^{(j-1)},s1)$,
- For σ_{V_i}: $p(\sigma_{V_i}^{(j)}|\sigma_{V_i}^{(j-1)}) \sim Gamma(c1,\sigma_{V_i}^{(j-1)}/c1)$,
- For a_{1i}: $p(a_{1i}^{(j)}|a_{1i}^{(j-1)},a_{2i}^{(j-1)},b_i^{(j-1)})$
  $\sim I_{(a_{1i}^{(j)})^2+(a_{2i}^{(j-1)})^2+(b_i^{(j-1)})^2\le 1} Uniform(\max(a_{1i}^{(j-1)}-c2,-1),\min(a_{1i}^{(j-1)}+c2,1))$,
- For a_{2i}: $p(a_{2i}^{(j)}|a_{1i}^{(j)},a_{2i}^{(j-1)},b_i^{(j-1)})$
  $\sim I_{(a_{1i}^{(j)})^2+(a_{2i}^{(j)})^2+(b_i^{(j-1)})^2\le 1} Uniform(\max(a_{2i}^{(j-1)}-c2,-1),\min(a_{2i}^{(j-1)}+c2,1))$,
- For b_i: $p(b_i^{(j)}|a_{1i}^{(j)},a_{2i}^{(j)},b_i^{(j-1)})$
  $\sim I_{(a_{1i}^{(j)})^2+(a_{2i}^{(j)})^2+(b_i^{(j)})^2\le 1} Uniform(\max(b_i^{(j-1)}-c2,0),\min(b_i^{(j-1)}+c2,1))$,
- For $\mu_\mu,\mu_{a_1},\mu_{a_2},\mu_b$: $p(\mu_*^{(j)}|\mu_*^{(j-1)}) \sim Normal(\mu_*^{(j-1)},s2)$,
- For $(\sigma_\mu,\sigma_{a_1},\sigma_{a_2},\sigma_b)$: $p(\sigma_*^{(j)}|\sigma_*^{(j-1)}) \sim Gamma(c3,\sigma_*^{(j-1)}/c3)$.

$s0,s1,c1,c2$, and $c3$ are tuning parameters controlling the step size.

The problem in implementing Gibbs and MH algorithms as above is posterior chain convergence. Because there are a large number of unknowns and correlations among them, it is hard to achieve convergence of posterior samples. Furthermore, updating one parameter at a time is not

only computationally inefficient, but it is also easily trapped in the wrong neighborhood of local extreme vaues. For example, because the distribution of $\ln V_{i,1:T}$ is heavily dependent on σ_{V_i}, if sampled σ_i values are small then volatility of sampled $\ln V_{i,1:T}$ again becomes small because posterior sampling of $\ln V_{i,1:T}$ is based on previous sampled σ_{V_i}. Once the chain is trapped in this circle, it is hard to get out when we update one parameter at a time. To solve this problem, we adjust a simple Gibbs sampling method by adopting the idea of simultaneous update of entire asset process $\ln V_{i,1:T}$ and its volatility σ_{V_i} (Liu and Sabatti (2000), Kou, Xie, and Liu (2005)). We move

$$((\ln V_{i,1:T} - \ln V_{i,1}), \sigma_i) \rightarrow (s_v(\ln V_{i,1:T} - \ln V_{i,1}), s\sigma_i), \tag{3.45}$$

where s_v is a scalar.

3.7.2 Estimation Procedure for Model 2: Correlated Intensity Model

For our correlated intensity model estimation, similar techniques from the previous subsection are adopted, with the following conditional distributions:

$$\begin{aligned} P_V(\lambda_{it}|\lambda_{i,t-1}, X_{1,1:t}, X_{2,1:t}, Y_{1:t}, \Theta_\lambda) &\sim Normal(\mu_{intensity}, \sigma_{intensity}), & (3.46)\\ P_{CDS}(\ln CDS_{it}|\lambda_{it}, \Theta_\lambda) &\sim Normal(l(\lambda_{it}, \mu_i', \sigma_i'), \sigma_{CDS}), & (3.47)\\ P_Y(\Delta Y_t|\Theta_\lambda) &\sim Normal(0, \Delta t). & (3.48) \end{aligned}$$

$\mu_{intensity}, \sigma_{intensity}$ are described in equation (3.32) and(3.33). $l()$ functions are described in equation (3.31).

As a proposal for λ, we use the gamma distribution, $Gamma(c4, \lambda_{it}^{j-1}/c4)$. We move λ and μ_{λ_i}' simultaneously:

$$(\lambda_{it}, \mu_i') \rightarrow (\lambda_{it} + 4s_\lambda, \mu_i' - s), \tag{3.49}$$

where s_λ is a scalar. This opposite direction of movement seems unreasonable in terms of intensity dynamic, which is a CIR specification. However, the survival probability, which is essential in

CDS pricing formula described in equation (3.30), is $\exp(A(i,t)\mu_i + B(i,t)\lambda_{it})$ where $A(i,t)$ and $B(i,t)$ are functions of μ_i' and σ_i'.

3.8 Simulation

In this section, we use simulated data to check the performance of the Bayesian estimation method adopted in this essay. We generate the 10 firms and one year daily data set. 2,000 posterior samples are drawn from the posterior distribution described in section 3.7. The burn-in sample period is 1,000 in each cases. The posterior sample convergence is checked with the potential scale reduction factor (psrf) (Gelman and Rubin (1992) and Brooks and Gelman (1998)) to ascertain the appropriateness of this posterior sample size and tuning parameters described in section 3.7. The psrf for all parameters are near 1 (appendix B.1).

3.8.1 Simulation Result for Model 1: Correlated Structural Model

Under the structural model, the daily asset value processes of 10 firms are generated. Based on equation (3.12) and equation (3.13), with one observed and one hidden common factor, we generate each firm's one year daily asset value process. Equity and 5-year CDS prices are derived from simulated asset value paths based on equations (3.14) and (3.15). We control parameters μ_i, σ_i and V_{i0} to make generated stock and CDS prices processes are close to the real market data. We diverge a_i and b_i as much as possible. Given all V_{i0}, to set an initial configuration from which the posterior samples of V_{it}, μ_i, σ_i are drawn, we adopt the iterative scheme as in Duffie and Wang (2007). Setting the initial configuration of posterior sampling based on such an iterative scheme reduces necessary sample size to make the posterior sample converge. The iterative scheme started at $V_{it}^{start\ iter} = (V_{i0}/S_{i0})S_{it}$. In the m^{th} iteration step, we calculate $\mu_i^{(m)}, \sigma_i^{(m)}$ and $V_{it}^{(m)}$ as below:

- For $\mu_i^{(m)}$: $\mu_i^{(m)} = \frac{\sum_{t=1}^{t=T}(V_{it}^{(m)} - V_{i,t-1}^{(m)})}{T-1}$
- For $\sigma_i^{(m)}$: $\sigma_i^{(m)} = \left(\sqrt{\frac{\sum_{t=1}^{t=T}(V_{it}^{(m)} - V_{i,t-1}^{(m)} - \mu_i^{(m)})^2}{T-1}}\right)/\sqrt{\Delta})$
- For $V_{it}^{(m)}$: Arg $\min_{V_{it}^{(m)}} \left(g_S(V_{it}^{(m)}, \sigma_i^{(m)}, T_D, D) - S_{it}\right)^2$

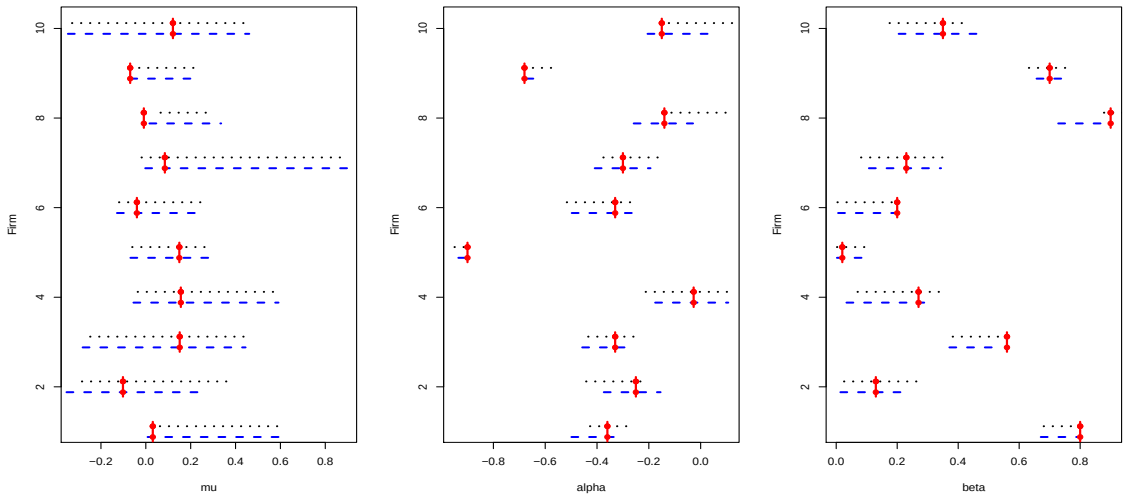

Figure 3.3: **Estimation Results of μ_i, a_i and b_i in the Simulated Structural Model.** These graphs are the posterior samples of parameters μ_i, a_i and b_i. Red vertical lines are true values. Black dotted lines are 95% posterior intervals using stock information only. Blue dashed lines are 95% posterior intervals using both CDS and stock information.

This iteration is continued until all V_{it}, μ_i and σ_i converge, and we set these converged values as the initial configurations of their posterior samples. Posterior samples of all other parameters (hyperparameters and random coefficients) are initialized from random points. A standard default barrier function in the Black-Cox model is adopted as in equation (3.19). We set $\gamma = 0.02, T_D = 10$ years. Recovery rates for all firms are set equal to 0.51 (which is known as the median recovery rate of historical data).

In estimating μ_i, a_i and b_i, as shown in Figure 3.3, all posterior distributions can correctly capture their true values with a similar variance in both cases: using only stock prices and using both stock and CDS prices information.

However, after adding the CDS information, significant improvements are seen in estimating volatility parameters σ_i. Table 3.2 presents a summary of the simulated posterior distribution using only stock information and using both CDS and stock information. In terms of both estimation accuracy and efficiency, additional CDS spreads information make significant improvements. The posterior mean while using both CDS and stock prices is closer to the true values, and the posterior standard deviation is much smaller when both CDS and stock prices are used.

Table 3.2: **Estimation Results for σ_i in the Simulated Structural Model.** This table summarize the posterior mean and posterior standard deviation of all 10 simulated asset volatilities.

	Firm	Firm1	Firm2	Firm3	Firm4	Firm5	Firm6	Firm7	Firm8	Firm9	Firm10
	True σ_i	0.4	0.25	0.42	0.29	0.36	0.13	0.49	0.18	0.33	0.52
With CDS	mean	0.4005	0.2500	0.4202	0.2900	0.3604	0.1304	0.4893	0.1808	0.3303	0.5204
and Stock	s.d.	0.0024	0.0008	0.0016	0.0098	0.0062	0.0038	0.0009	0.0086	0.0026	0.0013
With Stock	mean	0.3715	0.3385	0.4664	0.2893	0.3623	0.1297	0.4603	0.1803	0.3156	0.4900
	s.d.	0.0152	0.0113	0.0159	0.0181	0.0114	0.0043	0.0173	0.0051	0.0096	0.0202

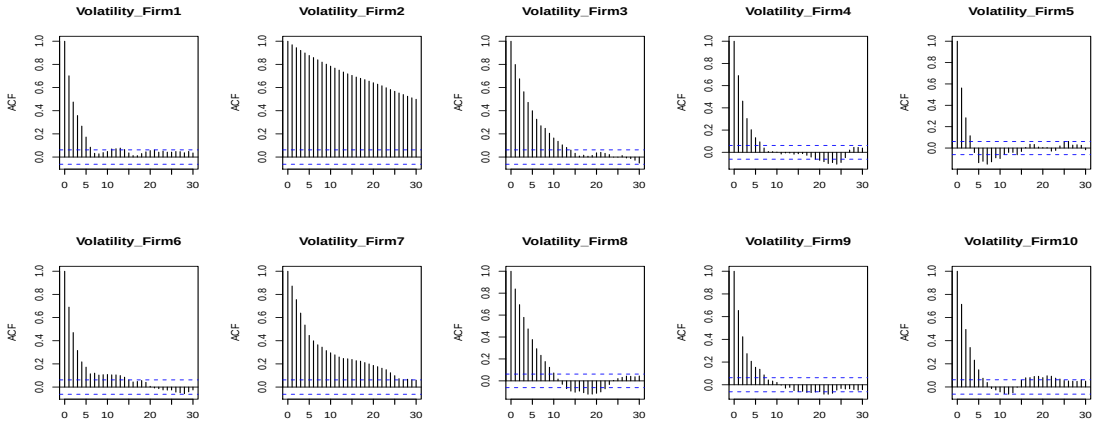

Figure 3.4: **ACFs of Volatility Posterior Sample Using Both CDS and Stock Data in the Simulated Structural Model.** These are ACFs of posterior samples of $\sigma_1, \ldots \sigma_{10}$ when we use both CDS and stock market information.

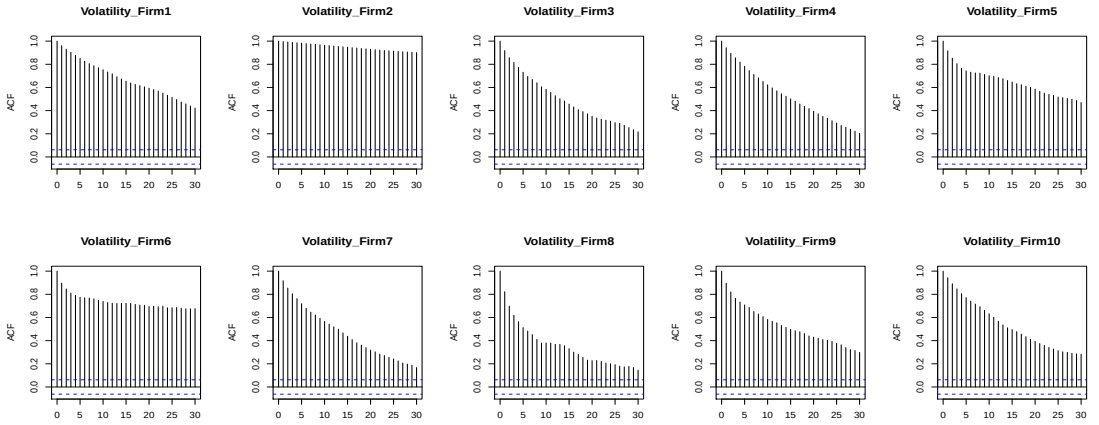

Figure 3.5: **ACFs of Volatility Posterior Sample Using Only Stock Data in the Simulated Structural Model.** These are ACFs of posterior samples of $\sigma_1, \ldots \sigma_{10}$ when we use only stock market information.

Furthermore, their posterior chain converges faster after adding CDS information as shown in Figures 3.4 and 3.5. Figure 3.4 shows the ACFs of posterior volatility sample when we use both CDS and stock information. Figure 3.5 shows the ACFs of posterior volatility sample when we use only stock information.

The hidden process and all the asset value paths are correctly estimated with or without CDS information. The results are shown in the appendix B.2.1.

3.8.2 Simulation Result for Model 2: Correlated Intensity Model

Under the intensity model, we generate the daily intensity processes of 10 firms for one year using equation (3.26) and equation (3.27) with one observed factor and one hidden factor. The CDS prices are then generated based on equation (3.28). We control parameters μ_i', σ_i' and λ_{i0} to make generated stock and CDS prices processes are close to the real market data. a_i' and b_i' are set equal to a_i and b_i values in the structural model simulation.

A similar iterative scheme is adopted to set an initial configuration from which posterior samples of λ_{it}, μ_i' and σ_i' are drawn as in the structural model. Given all intensities λ_{i0} at time 0, we set $(\lambda_{i0}/CDS_{i0})CDS_{it}$ as a starting point of the iterative scheme for λ_{it}. In the m^{th} iteration, $\mu_i'^{(m)}$ and $\sigma_i'^{(m)}$ are set to be the values that give the smallest $\left(l_{CDS}(\mu_i'^{(m)}, \sigma_i'^{(m)}, \lambda_{it}^{(m)}, T_{CDS}) - \ln CDS_{it}\right)$. This iteration is continued until all λ_{it}, μ_i' and σ_i' converge, and we set these converged values as the initial configurations of their posterior samples. All other parameters (hyperparameters and random coefficients) are initialized from random points. Recovery rates for all firms are set equal to 0.51.

Figure 3.6–3.9 display posterior distributions of a_{1i}', b_i'. The chain convergence is not as good as in the structural model, because only CDS information can be used in the intensity setting . However, by adopting a generalized Gibbs sampling scheme, we can estimate all unknowns correctly. The firm 9, which shows poorly fitted results, is an example of a firm which is significantly affected by both observed and unobserved common factors. For such a firm, estimates of factor coefficients are worse than for others. However, such cases do not happen in real data (will be shown in section 3.10). More simulation results are shown in the appendix B.2.2.

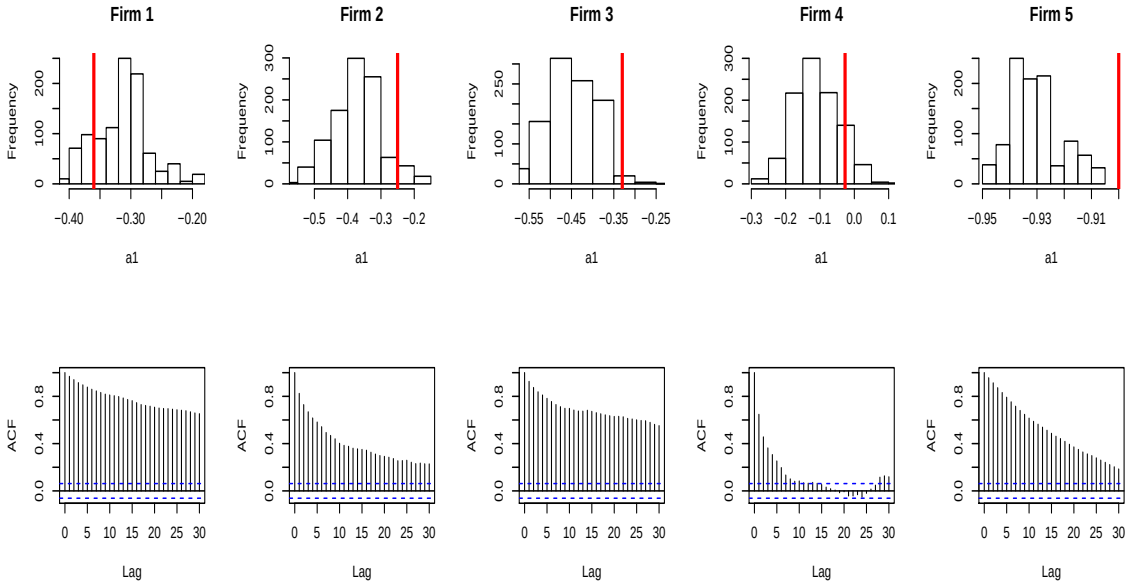

Figure 3.6: **Estimation of a_{1i} in the Simulated Intensity Model for Firms 1-5.** The top row of the graphs are the histograms of posterior sample of observed common factor effect a_1 for firm 1-5. The bottom row of the graphs are their ACFs. Solid red lines in the histograms are true values.

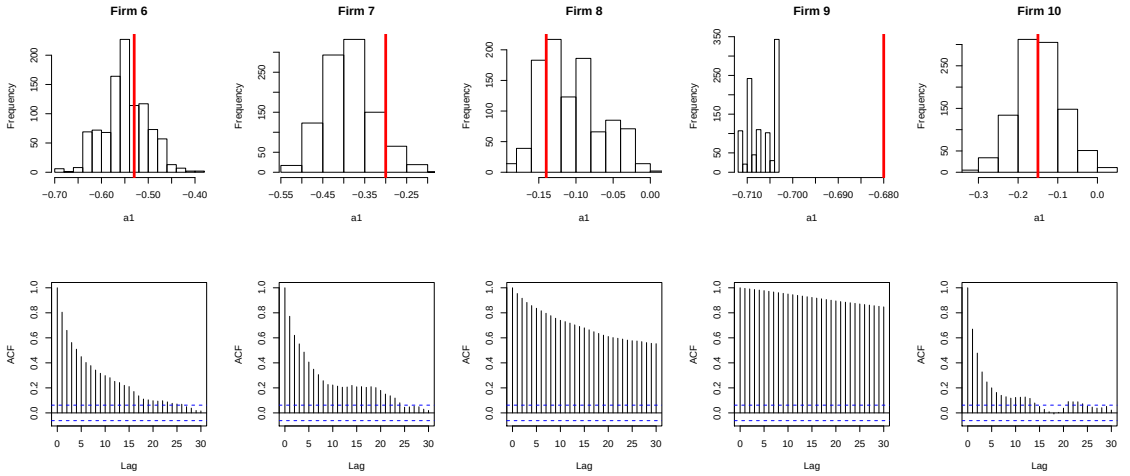

Figure 3.7: **Estimation of a_{1i} in the Simulated Intensity Model for Firms 6-10.** The top row of the graphs are the histograms of posterior sample of observed common factor effect a_1 for firm 6-10. The bottom row of the graphs are their ACFs. Solid red lines in the histograms are true values.

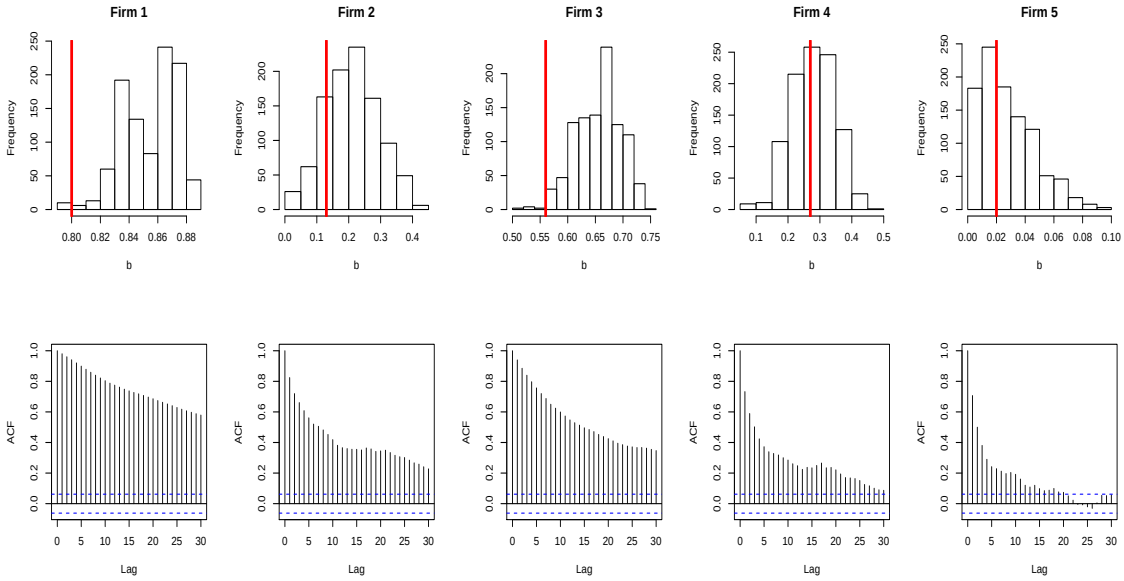

Figure 3.8: **Estimation of b_i in the Simulated Intensity Model for Firms 1-5.** The top row of the graphs are the histograms of posterior sample of observed common factor effect b_i for firm 1-5. The bottom row of the graphs are their ACFs. Solid red lines in the histograms are true values.

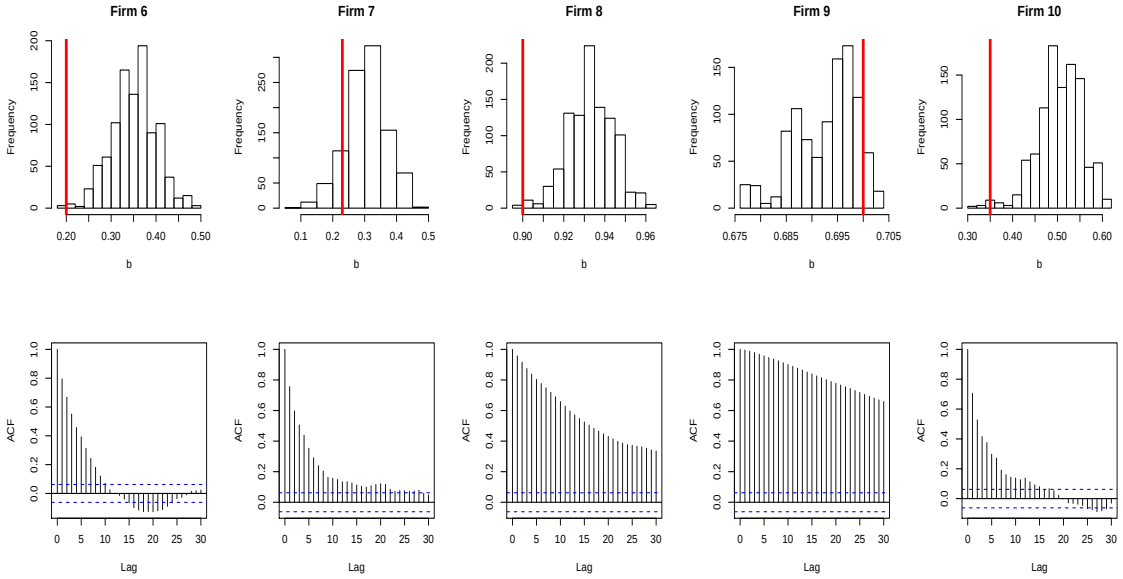

Figure 3.9: **Estimation of b_i in the Simulated Intensity Model for Firms 6-10.** The top row of the graphs are the histograms of posterior sample of observed common factor effect b_i for firm 6-10. The bottom row of the graphs are their ACFs. Solid red lines in the histograms are true values.

In this section, we check the appropriateness of the Bayesian estimation method described in section 3.7, because there are several unknowns in our model setting, including all the parameters, their hyperparameters, hidden process, and unknown state processes (asset or intensity). Through the simulated 10 firms' data sets, we found that the Bayesian estimation method we have adopted in this essay can correctly estimate all the unknowns. As multiple securities research, we also found that the additional CDS information enhances the estimation of volatility both in terms of efficiency and accuracy, in the structural model. In the next section, we will apply models and estimation methods to real data and explore empirical implications.

3.9 Data and Calibration

To apply our models and estimation methods to the real financial world, we first introduce the data set in this section. This section also discusses adjustment of unknowns, which are treated as exogenous variables because of the model identification problem.

3.9.1 Data Overview

We collect financial data on firms comprising CDX North American Investment Grade index series13
(CDX.NA.IG13). CDX.NA.IG is based on an equally weighted basket of CDS contracts for 125 North American firms having the same maturity investment grade corporate debt. The components of CDX are updated every 6 months in March and September. Thus, the first index CDX.NA.IG1 covers from September, 2003, and CDX.NA.IG13 covers up to December, 2009. Daily spreads of the 125 issuers' 5-year CDS comprising CDX.NA.IG13 are obtained from Bloomberg. The data encompasses the period from January 2, 2009, to December 31, 2009, and consists of mid-quote spreads. Among those 125 issuers, only 108 companies that have publicly traded stock in the U.S. and show no missing information on the balance sheet are used. Table 3.3 displays the deleted 17 companies and their deletion reasons. There are five deletion reasons. First, firms such as AT&T Mobility LLC are deleted because they are sub-divisions of one company. However, in the stock market, only company-level stocks are publicly traded. In the case of AT&T, both AT&T

Table 3.3: **17 Deleted Companies and Their Deletion Reasons**

Name	Ticker	Reason to be Deleted
AT& T Mobility LLC	T	Division of a company
Boeing Capital Corp.	BA	Division of a company
Boston Properties LP	BXP	Don't have publicly traded stock
Burlington Northern Santa Fe LLC	BRK	Don't have publicly traded stock
Canadian Natural Resources Ltd.	CNQCN	Canadian company
Capital One Bank, (USA) N.A.	COF	Division of a company
COX Communications Inc.	COXENT	Private company
DIRECTV Holdings LLC / DIRECTV Financing Co., Inc.	DTV	Division of a company
ERP Operating LP	EQR	Private company
General Electric Capital Corp.	GE	Division of a company
International Lease Finance Corp.	AIG	Don't have publicly traded stock
Motorola Solutions Inc.	MSI	Don't have publicly traded stock
National Rural Utilities Cooperative Finance Corp.	NRUC	Don't have publicly traded stock
News America Inc.	NWSA	Division of a company
Simon Property Group LP	SPG	Don't have publicly traded stock
Vornado Realty LP	VNO	Don't have publicly traded stock
Yum! Brands Inc.	YUM	Negative debt to equity ratio

Mobility and AT&T Corp. are included in the 125 issuers of CDX.NA.IG13, but only the stock for AT& T Corp. is available in the market. All other similar issuers are deleted from final data set. Second, firms such as Boston Properties LP have different types of business entities with public companies. Most of them are defined as private companies. For example, Boston Property Inc. has publicly traded stock information but such information does not exist for Boston Properties LP. Others are deleted because COX Communications Inc. and ERP Operating LP are private companies and Canadian Natural Resources Ltd. is a Canadian company. Yum! Brands Inc. is used for preliminary studies, which is described in section 3.3, but it is deleted from the final data set because the balance sheet information does not make sense. Summary statistics of the 108 companies included in empirical analysis are given in appendix B.3. Companies will now be referred to by the ticker symbol given in this table instead of the full name of the company.

Stock information is from CRSP and the balance sheet information is from Compustat. For a fixed interest rate, the average of one-year Treasury constant maturity rate of 0.0047 during 2009 from the Federal Reserve Bank is used.

3.9.2 Calibration of Default Barrier and Recovery Rate

As shown in equation (3.6), an additional parameter –recovery rate R– is needed to incorporate the CDS price into our model. However, it is well understood that separate identification of default intensities and recovery rate is not feasible (Das and Hanouna (2009)). Similarly, the unknown default barrier parameters, asset values, and recovery rate under the structural model setting are not identifiable. For this reason, most papers assume recovery rate as exogenously supplied values. Using the iterative scheme proposed by Vassalou and Xing (2004), Hull, Predescu, and White (2010) found default barrier parameters under the fixed recovery rate. They used the CDX.NA.IG index level at the mid-date of the data period and derived common default barrier parameters for all firms. However, in multiple-firm case, it is hard to assume that the default barriers are the same across all the different firms (Pan and Singleton (2008)). In practice, when Moody's periodically reports the average one year default rates, companies are grouped by their credit rating.

In order to reflect different default risks, default barrier parameters are calibrated. We set individual L_{0i} level in the default barrier function, which is described in equation (3.19). Another default barrier parameter γ is set to be 0.02 as in Lando (2004)[10]. We fix the recovery rate $R = 0.51$, which is the median recovery rate in historical data.

The m^{th} iteration step, brings us from an estimate σ_i^m and L_{0i}^m to an improved estimate σ_i^{m+1} and L_{0i}^{m+1}.

- For all i and t, find V_{it}^m which minimizes the value $g_S(V_{it}^m, \sigma_i^m, T_D, D) - \ln S_{it}$, where $g_S()$ is given in equation (3.20).

- For all i, estimate σ_i^{m+1} by thinking of $V_{i1}^m, ..., V_{iT}^m$ as geometric Brownian motion, i.e., let

$$\sigma_i^{m+1} = s.d.(\ln V_{it}^m - \ln V_{it-1}^m)$$

- For all i, find the value L_{0i}^{m+1} which minimizes $\sum_{t=1}^{T}(g_{CDS}(V_{it}^m, \sigma_i^m, T_{CDS}) - \ln CDS_{it})^2$ with

[10] Previous research on structural credit risk models usually assume γ takes a value between 0 to 0.04. We found that, within this range, the fixed γ value does not affect significantly the determination of L_{0i} levels.

a constraint that $V_{it} > L_{i0}\exp(-\gamma(T_D - t))$ for t=1,...T, where $g_{CDS}()$ is given in equation (3.22) and depends on the value of L_{0i}.

Using this updated value, the procedure continues with σ_i^{m+1} in place of σ_i^m and with L_{0i}^{m+1} in place of L_{0i}^m until they all converge.

3.10 Empirical Results

3.10.1 Empirical Results for Model 1: Correlated Structural Model

This subsection discusses the empirical results using stock and CDS market information under our model 1: Correlated structural model. For the initial configurations of posterior sampling are set as in the simulation study. The sum of liability and equity value in December 2008 is set as a first value of asset of each firm V_{i0}.

The Role of Additional CDS Information: Estimation with Only Stock versus Estimation with Both CDS and Stock

The usual structural model is estimated only with stock information, so before we show the estimated correlation structure, we discuss the difference additional CDS information makes in our estimation. We first fit the model only with stock information and then add CDS spreads information. The posterior sample are drawn 2,000 times with a 1,000 burn-in period, in each fitting.

Figures 3.10 and 3.11 show the difference between before and after using the CDS information. Figure 3.10 shows the percentage increase, ξ, in posterior mean of sampled common factors coefficients a_{1i}, a_{2i} and b_i and volatility parameter σ_i when we add CDS information in the estimation process. ξ is defined as follows:

$$\xi = \frac{(\textit{Posterior mean when we use both CDS and S}) - (\textit{Posterior mean when we use only S})}{\textit{Posterior mean when we use only S}}.$$

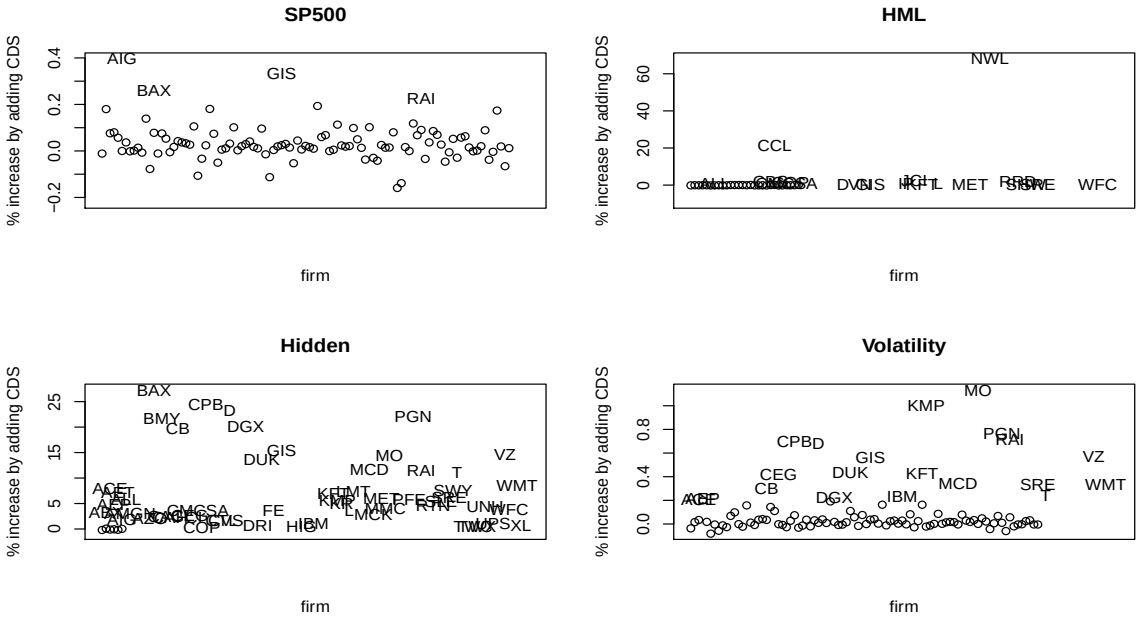

Figure 3.10: **% Increase of Posterior Mean of Parameters When Additional CDS Prices Are Used in Estimation.** The four graphs provide the % increase of posterior mean of parameters when additional CDS prices are used in estimation, which is ξ. The top left panel is for SP500 coefficient a_1. The top right panel is for HML coefficient a_2. The bottom left panel is for hidden factor coefficient b. The bottom right panel is for asset process volatility parameter σ_i.

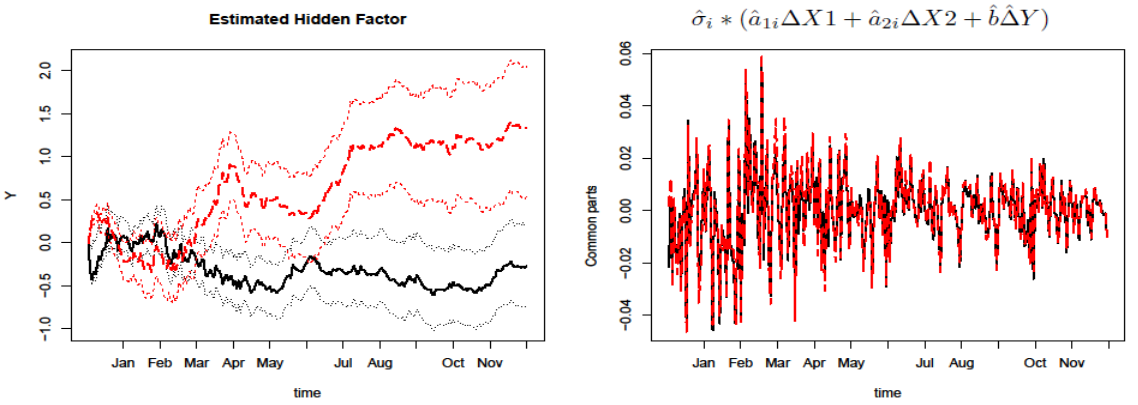

Figure 3.11: **Estimated Hidden Path and Mean Path of Common Factor Effect in the Structural Model.** The left graph is the estimated hidden path. Black solid line is the posterior mean of sampled hidden path using both CDS and stock prices and black dotted lines are their 95% posterior interval. Red long-dashed line is the posterior mean of sampled hidden path using only stock price and red dashed lines are their 95% posterior interval. The right graph is the estimated $\hat{\sigma}_i * (\hat{a}_{1i}\Delta X1 + \hat{a}_{2i}\Delta X2 + \hat{b}\hat{\Delta}Y)$. All $\hat{\sigma}_i, \hat{a}_{1i}, \hat{a}_{2i}, \hat{b}$ and $\hat{Y}$ values are means of each posterior sample. Black solid line is the means of estimated path using both CDS and stock prices. Red long-dashed line is the mean of estimated path using only stock price.

The left panel of Figure 3.11 shows the difference in sampled hidden factor paths. At first glance, it seems that there is a large difference in the estimated hidden path between the two. Two hidden factors seem to move in the opposite direction. For example, from February to April, hidden factor estimated with only stock prices information ($\hat{Y}_{1:T}^{(S)}$) is increasing, while hidden factor estimated with both stock and CDS prices information($\hat{Y}_{1:T}^{(CDS+S)}$) is decreasing. However, in the light of the fact that CDS can be used as *Early Warning Indicator* on the financial market, we find that $\hat{Y}_{1:T}^{(CDS+S)}$ acts earlier than $\hat{Y}_{1:T}^{(S)}$. If we push $\hat{Y}_{1:T}^{(CDS+S)}$ a little bit to the right, then both hidden factor estimates move in the same direction (even though there is still scale difference). In terms of the scale of hidden factor movement, $\hat{Y}_{1:T}^{(S)}$ is larger. It fluctuates between -0.5 to 2 while $\hat{Y}_{1:T}^{(CDS+S)}$ fluctuates between -0.6 to 0.1. This scale difference might lead the result that estimated hidden factor coefficients are generally larger when we use both stock and CDS prices.

$\hat{Y}_{1:T}^{(S)}$ is similar to the movement of HML, but $\hat{Y}_{1:T}^{(CDS+S)}$ shows opposite fluctuations of HML. However, as shown in Figure 3.10, there are large differences in the estimated coefficients a_{2i} and b_i too; we need to compare two results taking into consideration all estimation results together. In terms of overall comparison of the common factor effects, the right panel of Figure 3.11 depicts the mean of the common movement in asset dynamic which is $\sigma_i(a_{1i}\Delta X1 + a_{2i}\Delta X2 + b_i\Delta Y)$. Even though estimated b_i and a_{2i} are generally larger when we use CDS and stock prices together, $\sigma_i(a_{2i}\Delta X2 + b_i\Delta Y)$ remains similar before and after adding CDS information because $\hat{Y}_{1:T}^{(CDS+S)}$ moves in the opposite direction of HML movement but $\hat{Y}_{1:T}^{(S)}$ moves in the same direction of HML. Moreover, as shown in Figure 3.10, the most significant factor SP500 coefficients a_{1i} change less than the others, while the HML coefficient a_{2i} changes the most (the mean of all sampled $|a_{1i}|$ is 0.733, but the mean of all sampled $|a_{2i}|$ and b_i are 0.119 and 0.122, respectively, when only the stock information is used in estimation). Because there is less change in the more significant coefficient and estimated hidden factor paths and estimated b_i and a_{2i} all change together after adding CDS prices, the overall common factor effect remains similar. In terms of common factor movement estimation, additional CDS information does not offer a significant difference.

However, CDS information offers considerable difference in asset volatility estimation. As demonstrated in Figure 3.10, volatilities are generally estimated with larger values after adding

the CDS information. It is acknowledged that previous structural credit risk models had a defect in capturing long left tail behavior of asset value distribution. In other words, historical (real) asset values include extremely small values, even those are very rare, while asset value distribution estimation through the typical structural model cannot capture those rarely small cases. By offering larger volatility estimates, incorporating CDS information is expected to resolve this shortcoming to some degree. Moreover, adding CDS information in estimation gives larger common factor coefficients a_{1i}, a_{2i} and b_i and this lead larger correlation coefficients because correlation coefficients are determined by sum of the cross-products of common factor coefficients as shown in equation (3.25). As we show in Chapter 2 (Figure 2.5), larger estimated correlation coefficients improve the estimation of the tail behaviors of default distribution.

Another benefit of adding CDS information is in estimation efficiency. Adding CDS information improves the efficiency of the estimates. Figure 3.12 is a boxplot of the standard deviation difference of the posterior distribution of 108 firms' $a_{1i}, a_{2i}, b_i, \mu_i$, and σ_i between before and after adding CDS. Except for the hidden factor effect (marked with b_i), the posterior sample based on both CDS and stock information has a smaller posterior standard deviation. Asset price volatility estimation improved especially in terms of the estimation efficiency by adding CDS information. Moreover, similar to the simulation study in section 3.8, we find that algorithms converge faster in cases using CDS data.

Because estimation performance is improved by adding CDS, all the results given in the next two subsections reflect when both CDS and equity price are used for estimation.

Estimated Parameters and Estimated Correlation in the Correlated Structural Model

Figure 3.13 shows scatter plots of 108 firms' posterior mean for each pair of parameters. Widely spread factor coefficients show the necessity for relaxing the assumption on equal factor effects. The firms with larger mean values have larger volatility values except few outliers. Extremely large volatilities compared to their mean level is due to one or both of the financial markets (CDS and/or stock markets) conditions. The firms having large volatilities such as AIG, HIG, and XL suffered higher CDS spreads than other firms (their highest CDS spread during the year were

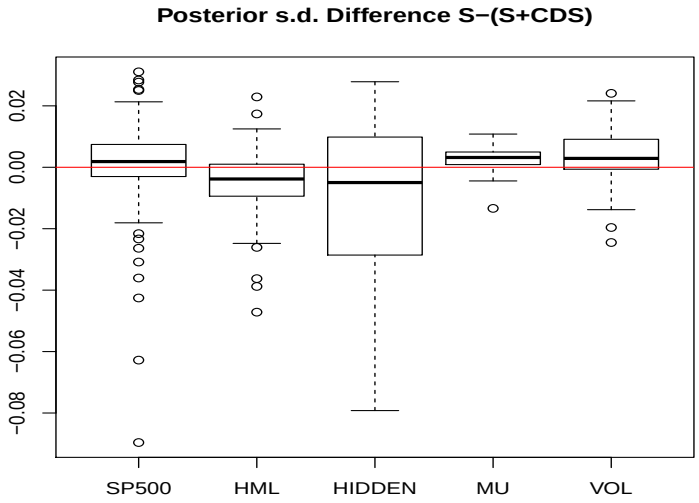

Figure 3.12: **Boxplots of Posterior Standard Deviation Difference between With and Without Using CDS Prices.** This graph shows the difference of the posterior standard deviation between before and after adding CDS prices in estimating for (from left) SP500 coefficient, HML coefficient, hidden factor coefficient, mean level of asset process, and volatility level of asset process. The four boxplots provide *(posterior standard deviation when we use only S) − (posterior standard deviation when we use both CDS and S).*

3758.987bp, 1122.27bp and 989.724bp, respectively, while the mean of highest CDS spreads for other firms are only 264.3929bp). In other words, the market risk assessment about these firms are higher than others and this leads to higher volatility estimates. However, highest CDS price of the firms such as CTL, WFC, and TWC are just 146.551bp, 304.125bp and 367.525bp, but their estimated volatilities are very high. These firms' high volatilities are due to the sudden increase or decrease in the equity price. In these firms, moreover, there were sudden changes in the number of outstanding, and this makes their asset volatility large.

The effect of SP500 is the most significant. The effects of SP500 are gathered near 0.8, which is about 4 times larger value than other common effect coefficients. The SP500 coefficient of CTL is 0.324 while that of HON is 0.882. In most firms, the sign of coefficients of HML is negative, and the unexpected abnormal negative signs of coefficients of HML occur because of the correlation among the factors. The effects of the hidden factor are relatively smaller than the observed factor effects but they are widely spread than other common effect coefficients.

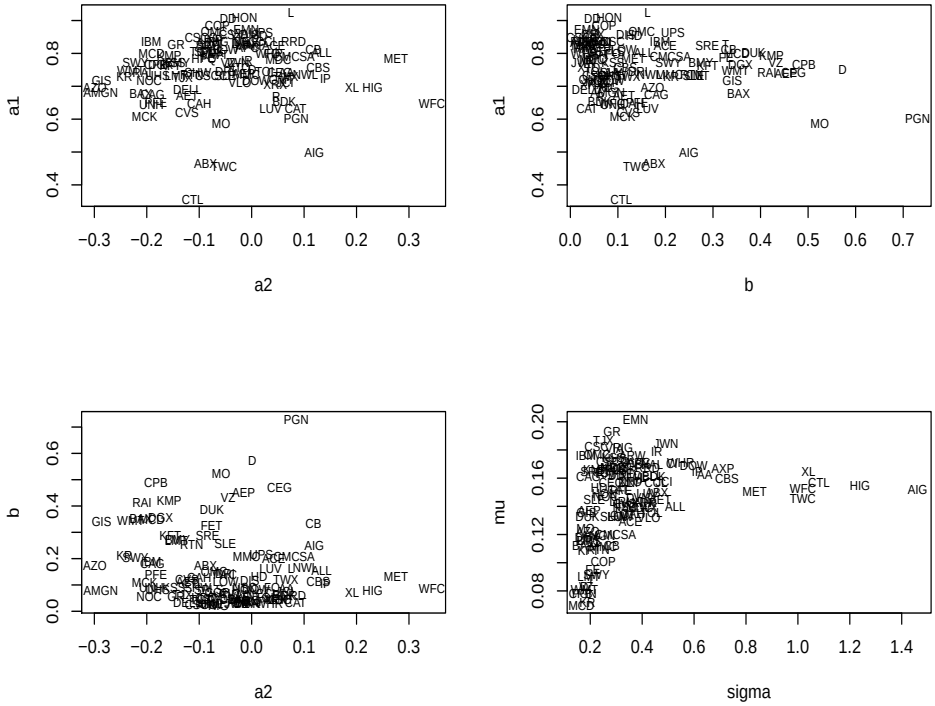

Figure 3.13: **Posterior Mean of Parameters in the Structural Model.** These are pairwise scatter plots of posterior means of parameters. The top left plot is SP500 coefficients versus HML coefficients. The top right plot is SP500 coefficients versus hidden factor coefficients. The bottom left plot is hidden factor coefficients versus HML coefficients. The bottom right plot is mean levels versus volatility levels of asset process.

More empirical results (hyperparameter and correlation coefficient estimates) are given in appendix B.4.1. We will be back to estimated hyperparameter values in comparison with intensity model in Table 3.5. Generally, assets of firms are highly correlated with each other. However, the AIG, ABX, CTL, TWC, and WFC firms have relatively low correlation coefficients. These are the firms with extremely small a_{1i} values and large and positive a_{2i}, as shown in Figure 3.13.

The Role of Hidden Factor in the Correlated Structural Model

In order to see the role of the hidden factor in the correlated structural model, we summarize results for each industry in Table 3.4. Because SP500 return is the most significant covariate

Table 3.4: **Industry Performance in S&P500 Return and Mean Common Factor Effect in Each Industry.** We summarized return performance and estimated results in each industry. (%) SP is proportion of firms which belongs to each industry in components of S&P500 index. (%) SP Return is percentage of return increase in each industry during 2009. Industry performance in S&P500 return increase during 2009 is summarized with (%) SP Share. (%) CDX is weight of each industry in CDX.NA.IG13. $\bar{a}_1, \bar{a}_2$ and $\bar{b}$ refer to average of the posterior mean of each coefficient in each industry.

Sector	(%)SP	(%)SP Return	(%)SP Share	(%) CDX	$\bar{a}_1$	$\bar{a}_2$	$\bar{b}$	$\bar{a}_1^2+\bar{a}_2^2+\bar{b}^2$
Info-Tech	15.27	59.92	38.985	7.4	0.78	-0.09	0.06	0.62
Cons-Disc	8.4	38.76	13.872	24.1	0.77	-0.02	0.11	0.60
Health Care	14.79	17.07	10.757	10.2	0.69	-0.17	0.16	0.53
Financials	13.29	14.80	8.38	10.2	0.75	0.14	0.16	0.61
Industrials	11.08	17.27	8.153	14.8	0.78	-0.03	0.10	0.61
Energy	13.34	11.29	6.417	7.4	0.81	-0.05	0.09	0.66
Cons-Stap	12.88	11.20	6.146	10.2	0.71	-0.18	0.31	0.64
Materials	2.93	45.23	5.646	6.5	0.74	0.00	0.09	0.55
Utility	4.19	6.80	1.214	6.5	0.75	-0.02	0.46	0.78
Telecom	3.83	2.63	0.429	2.8	0.65	-0.07	0.29	0.51

explaining asset levels of firms, industry performance in S&P500 return is also summarized in Table 3.4. The information-technology industry was up an impressive 59.92% during 2009 and its increase comprise a 38.985% share in the total S&P500 return increase during 2009 with consideration of its weight in S&P500 index. On the other hand, the utility industry was up only 6.8%, and it comprise only 1.2% in 2009 S&P500 return increases. Since the general movement of the S&P500 return during 2009 may have depended more on return movements of the information-technology firms rather than that of the utility firms, including the hidden factor is expected to incorporate small return movements, which are not captured well with overall macroeconomic variables. The results in Table 3.4, which shows larger hidden factor coefficients $\bar{b}$ in industries with smaller (%) SP Share, support this expectation.

There might be an argument that SMB (which is described in section 3.3) can incorporate small return movements into the analysis without complicated hidden factor estimation. However, the hidden factor can give us more data-oriented common movement. If the target portfolio is different, small return movement not captured by SP500 might be different. Instead of simply adding the observed SMB factor to explain the un-captured common movement to the SP500, by adding the hidden factor, we can get a more data-dependent common factor process. In turn, we can make the model be able to track un-captured common movement better and, moreover, we can

also understand the asset correlation structure better. Hidden factor analysis might be able to give more accurate correlation structure if target portfolio is smaller and more specified.

Furthermore, the hidden factor gives more information than the small return movement. It plays a role in adjusting two different weights in S&P500 companies and our target sample CDX companies. Table 3.4 shows that the estimated hidden factor coefficient $\bar{b}$ is not simply increasing as (%) SP Share is decreasing. By comparing the information technology, energy, materials and utility industries (which are the industries comprising CDX with similar proportion), we can find that as $\{$(%) SP Share - (%) CDX$\}$ becomes smaller (towards the negative), the hidden factor coefficient $\bar{b}$ increases.

There have been considerable works and discussions related to the issues on finding good covariate sources to explain financial market or portfolio returns. In this sense, hidden factor analysis can give another direction for those studies.

The last column in Table 3.4, $\bar{a}_{1i}^2 + \bar{a}_{2i}^2 + \bar{b}_i^2$ values are given. This value is the proportion of the asset volatility explained by the common factors. We call it commonality. The commonalities are generally high, but by comparing commonalities of the health care, financials and consumer staples industries which have similar number of firms in the sample, the firms are more correlated when the return is low. This is also true for a comparison of information-technology, materials, energy, and utility industries. In light of the fact that, generally, during a period of distress, defaults are more highly correlated than in stable periods, these results show a higher correlation in industries that are suffering from a downturn, which seems reasonable.

The Role of Flexibility in Coefficients in the Correlated Structural Model

The other extension we apply in this essay is relaxation of equal common factor coefficients across the firms. Figure 3.14 shows the posterior sample SP500 coefficients a_i and, for each firm, its 95% posterior interval. There is a significant amount of variation in a_1, a_2 and b. For firms such as ABX, AIG, CTL and TWC, asset movements are quite insensitive to observed market return than for the firms DD, HON and L. This suggests that Duffie and Saita (2009)'s approach to let all firms have the same sensitivity is not supported by the data.

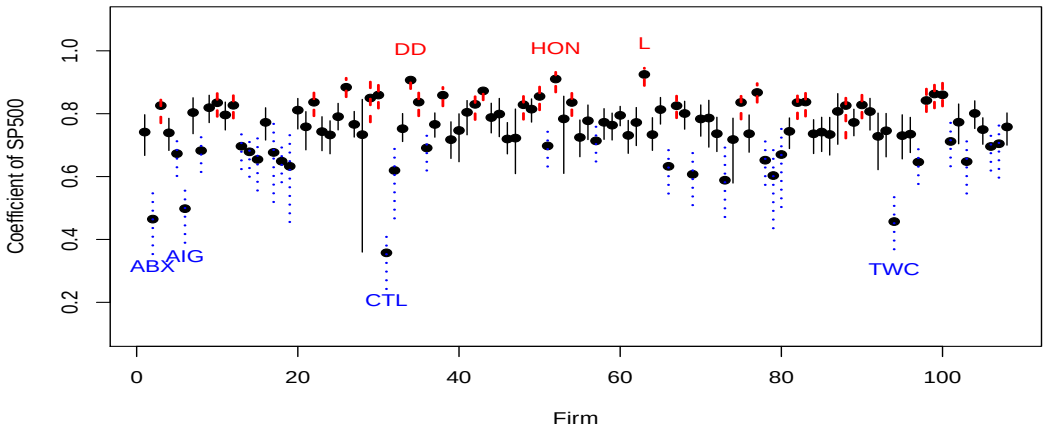

Figure 3.14: **Estimated SP500 Effects and Their 95% Posterior Interval(PI) in the Correlated Structural Model.** This figure shows all estimated a_{1i} and their 95% PI. Dotted blue lines refer to 95% PI of companies with relatively small market return effect (lower 25%). Dashed red lines refer to 95% PI of companies with relatively large market return effect (upper 25%) The rest of the companies' 95% PIs are marked with solid black lines.

3.10.2 Empirical Results for Model 2: Correlated Intensity Model

This subsection shows the empirical results using CDS spreads under Model 2: Correlated intensity model. As an initial configuration of posterior sampling for λ_{i0}, we set as below:

$$\lambda_{i0}^{initial} = -\ln(1-\Pr(\tau_i < 5))/5, \tag{3.50}$$

where

$$\Pr(\tau_i < 5) = 1 - \left(\frac{1-R_i}{(1-R_i)+(CDS_{i1}/4)}\right)^{20}. \tag{3.51}$$

CDS_{i1} is the annualized CDS spread on Jan 2, 2009. $\Pr(\tau_i < 5)$ is the default probability of the i^{th} firm during 5 years, and it is approximated as in equation (3.51) because the survival probability q_i during every quarter is determined as follows:

$$q_i CDS_{i1}/4 = (1-q_i)(1-R_i). \tag{3.52}$$

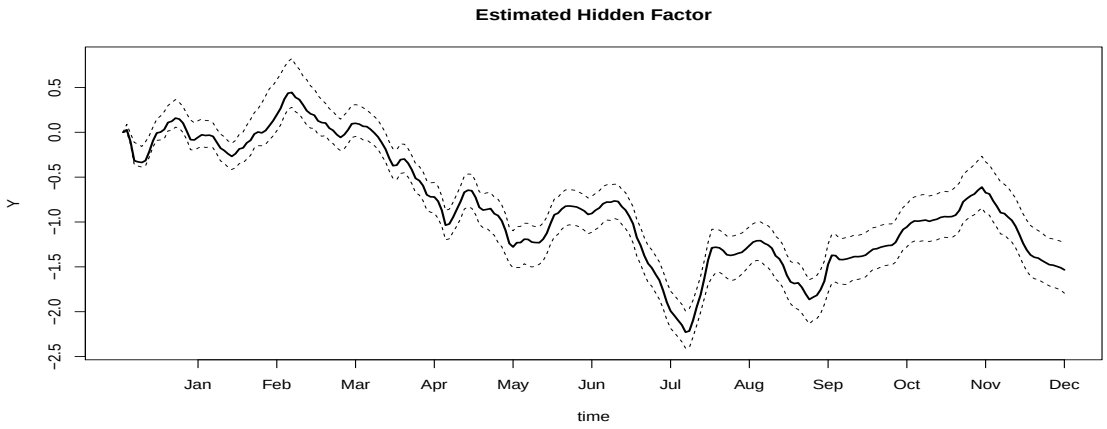

Figure 3.15: **Estimated Hidden Path in the Intensity Model.** Black solid line is the posterior mean of sampled hidden path and black dotted lines are its 95% posterior interval.

Similar iterative scheme then are applied as in the simulation study to set the initial configuration of posterior sampling for μ_i' and σ_i' and all λ_{it}. Initializations of other parameters are given also in section 3.8. Posterior samples are drawn 2000 times with a 1000 burn-in period.

Estimated Parameters and Correlation in the Intensity Model

Figure 3.15 depicts the posterior mean of the hidden path and their 95% posterior intervals. The estimated hidden path has an opposite fluctuation with that of the structural model because b is assumed to be positive. During second half of 2009, which experienced a high SP500 period and more stable HML period, the estimated hidden values were relatively lower than the values in the first half of 2009.

Figure 3.16 shows the estimated parameters. There is significant amount of variation in a_1, a_2, and b. The most obvious outliers are AIG and CAH. AIG, which had the largest CDS value during 2009, has extremely large volatility and mean level. Even though there was no large CDS value movement as in AIG, there were two comparable peak points in CDS price fluctuation of CAH during 2009. Other firm's CDS prices increased sharply during January and February, after which the prices stabilized; however, the CDS price of CAH increased again in December 2009 as much as in January and February of 2009.

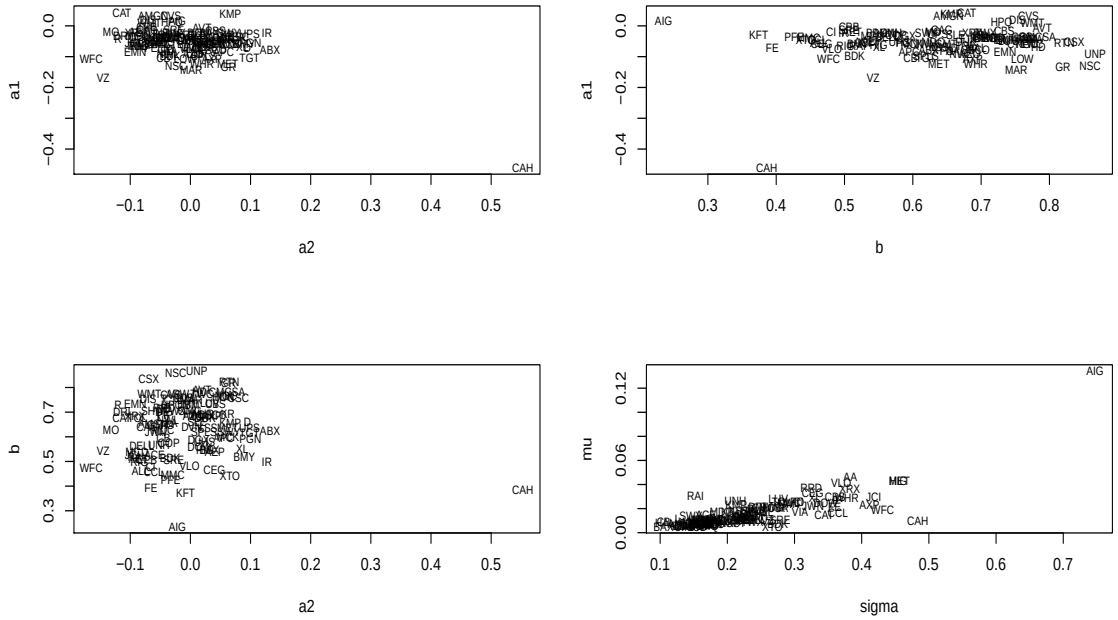

Figure 3.16: **Posterior Means of Parameters in the Intensity Model.** These are pairwise scatter plots of posterior mean of parameters. The top left plot is SP500 coefficients versus HML coefficients. The top right plot is SP500 coefficients versus hidden factor coefficients. The bottom left plot is hidden factor coefficients versus HML coefficients. The bottom right plot is mean levels versus volatility levels of intensity process.

In contrast with the structural models, the effects of the observed factors are very small compared to the effects of the hidden factor. The most significant common factor is hidden and the spread of b_i is also the largest. Only in the CAH firm are a_{1i}, a_{2i} and b_i comparable in absolute scale. This might be because the selection of observed factors depends on previous asset pricing or stock return predictability research. In explaining intensity movement, even SP500 has only marginal effect. In fact, this result corresponds to results of Duffie, Eckner, Horel, and Saita (2009), who used direct modeling without assumption of the dynamic of the intensities. These results lead us to infer that adopting the hidden factor is more necessary in the intensity model than in the structural model unless we can find other general economic variables substituting SP500.

More empirical results (hyperparameter and correlation coefficient estimates) are given in the Appendix B.4.2. Convergence of the posterior sample of hyperparameters governing the distribu-

tion of observed factor (SP500 and HML) effect is not as good as in the structural model because, as mentioned before, the effects of observed factors are marginal compared to the hidden factor effect. Hyperparameters governing the distribution of hidden factor effect and the log-transformed mean level of intensity process converge well. AIG has the smallest intensity correlation with others as it does in the structural model.

The Role of Hidden Factor in the Correlated Intensity Model

As mentioned earlier, the distribution assumptions on each common factor effect are adopted and hyperparameters governing those prior distributions are estimated from the data. By comparing these estimated hyperparameters, comparison of overall common factor effects between two different models is possible.

Table 3.5 shows the posterior mean of the sampled hyperparameters. We also calculate the conditional mean and conditional standard deviation of truncated normal hyper as follows: Suppose $x \sim N(\mu_x, \sigma_x)$ and lies within an interval (m, m') then

$$E(x|m<x<m') = \mu_x + \sigma_x \frac{\phi(\frac{m-\mu_x}{\sigma_x}) - \phi(\frac{m'-\mu_x}{\sigma_x})}{\Phi(\frac{m'-\mu_x}{\sigma_x}) - \Phi(\frac{m-\mu_x}{\sigma_x})}, \tag{3.53}$$

$$sd(x|m<x<m') = \sigma_x \left(1 + \frac{\frac{m-\mu_x}{\sigma_x}\phi(\frac{m-\mu_x}{\sigma_x}) - \frac{m'-\mu_x}{\sigma_x}\phi(\frac{m'-\mu_x}{\sigma_x})}{\Phi(\frac{m'-\mu_x}{\sigma_x}) - \Phi(\frac{m-\mu_x}{\sigma_x})} - \left(\frac{\phi(\frac{m-\mu_x}{\sigma_x}) - \phi(\frac{m'-\mu_x}{\sigma_x})}{\Phi(\frac{m'-\mu_x}{\sigma_x}) - \Phi(\frac{m-\mu_x}{\sigma_x})} \right)^2 \right) \tag{3.54}$$

Even though a direct comparison of the effect of each factor between two different models is not plausible, we can still compare the effect of each factor in relative scale. Under the structural model, SP500 has the largest effect on asset correlation and the effect of the hidden factor is only about $1/3$ of the effect of SP500. In contrast, in the intensity model, the hidden factor has the largest effect on the intensity correlation, while effects of other observed factors are marginal compared to the effect of the hidden factor. This result shows that the hidden factor plays an important role in both models, but in terms of relative magnitude, the hidden factor is more necessary in the intensity model.

Incorporating a hidden factor is necessary in the structural and intensity models because it can

Table 3.5: **Comparison Posterior Mean and S.D. of Hyperparameters between Correlated Structural Model and Correlated Intensity Model.** This table shows the posterior mean of hyperparameters μ_x and σ_x(in parenthesis). Because prior for a_1, a_2 and b are truncated normal distributions, conditional means and standard deviation are also calculated (for a_1 and a_2, truncated points $m = -1$ and $m' = 1$ and for b, $m = 0$ and $m' = 1$.)

Unconditional	Model	$\mu_{a_1}(\sigma_{a_1})$	$\mu_{a_2}(\sigma_{a_2})$	$\mu_b(\sigma_b)$
$\mu = E(x)$	Inten	-0.055(0.078)	-0.005(0.107)	0.635(0.128)
$\sigma = sd(x)$	Struc	0.754(0.106)	-0.044(0.131)	-1.914(0.589)
Conditional	Model	$Cond\mu_{a_1}(Cond\sigma_{a_1})$	$Cond\mu_{a_2}(Cond\sigma_{a_2})$	$Cond\mu_b(Cond\sigma_b)$
$Cond\mu = E(x\|m < x < m')$	Inten	-0.055(0.078)	-0.005(0.107)	0.634(0.126)
$Cond\sigma = sd(x\|m < x < m')$	Struc	0.751(0.102)	-0.044(0.131)	0.162(0.149)

explain asset and intensity correlation, which is not captured by the observed common factors. However, its role under each model is different. In the structural model, as shown in the previous section, the hidden factor plays a more important role in explaining the correlation of the firms with low return and less contribution to the S&P500 return. In terms of capturing diversified correlation structure among firms' assets, the hidden factor contributes in the structural model. However, the hidden factor is certainly necessary in capturing basic intensity correlation because, in the intensity model, observed factors cannot capture the intensity correlation as much as they did in the structural model.

The Role of Flexibility of Coefficients in the Correlated Intensity Model

There is a significant amount of variation in a_1, a_2, and b. Figure 3.17 shows the posterior sample of hidden factor coefficients b_i and their 95% posterior intervals. As in our correlated structural model, the generalization that all firms have the same sensitivity is not supported by the data.

3.11 Pricing CDS Index Tranches Prices

In order to ascertain whether our correlation estimates are in line with empirical finding, this section introduces the CDX.NA.IG13 tranches. We test how well the models fit market data on the prices of CDX tranche and check the model validation with this out-of-sample data. To value a CDX tranche, it is necessary to develop a model of the joint probability of default of the companies

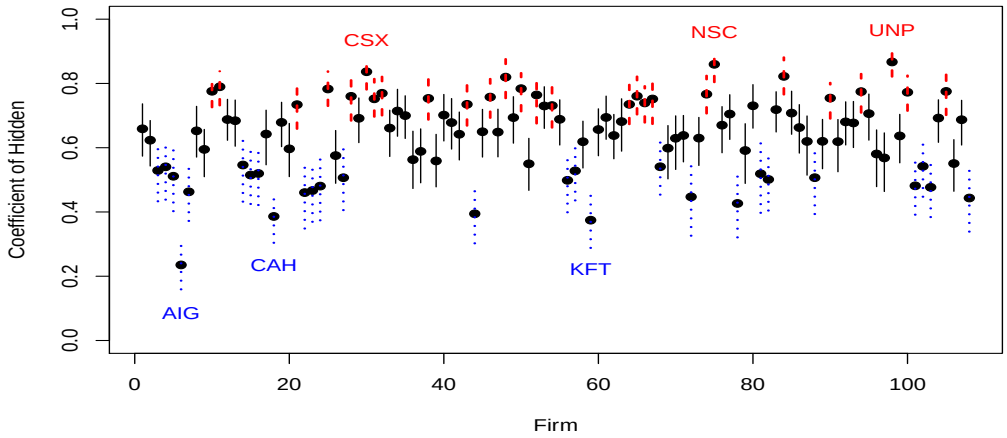

Figure 3.17: **Estimated Hidden Factor Effects and Their 95% Posterior Interval (PI) in the Correlated Structural Model.** This figure shows all estimated b_i and their 95% PI. Dotted blue lines refer to 95% PI of companies with relatively small market return effect (lower 25%). Dashed red lines refer to 95% PI of companies with relatively large market return effect (upper 25%). The rest of the companies' 95% PIs are marked with solid black lines.

in the underlying portfolio, so the market price of CDX tranche could be used to test the plausibility of our estimated results. As an example of typical financial derivatives related to multi-name defaults, CDX tranche pricing shows the potential uses of our models for correlated defaults.

Overview on CDS Index Tranches

In Section 3.9, we introduced the CDS index. CDS index tranches are also issued, as a synthetic CDO, with each tied to a specific CDS index. It is easiest to think of CDS index tranches as a basket of CDS, which provides the layer protection. CDX tranches with maturity T_b are contracts involving protection buyer and seller. There are two default loss levels $a_H\%$ and $a_L\%$ for a reference portfolio. The protection seller absorbs the $a_L\%$ to $a_H\%$ of the total loss on the portfolio of the buyer. In turn, protection buyer must make periodic payments with annualized rate $Spr_{L,U}$. CDX tranche pricing was introduced in Torresetti, Brigo, and Pallavicini (2007) and Longstaff and Rajan (2008).

The loss of a_L%-a_H% tranche $L_t^{L,U}$ is as follows:

$$L_t^{L,U} = \frac{1}{a_H - a_L}\max(\min(Loss_t, a_H) - a_L, 0), \quad \text{where } Loss_t = \frac{\sum_{i=1}^{N}(1-R_i)1_{D_{it}=1}}{N}, \tag{3.55}$$

where D_{it} is a default indicator of i^{th} firm by time t and R_i is the recovery rate of firm i; N is the total number of firms.

The protection seller covers the loss from a_L% to a_H% in each discretized time (t_{i-1}, t_i), so the present value of the protection seller's payment is

$$E\left(\sum_{i=1}^{M} DFD(0,t_i)(L_i^{L,U} - L_{i-1}^{L,U})\right). \tag{3.56}$$

The protection buyer pays with annualized rate $Spr_{L,U}$ on the reduced notional principal due to defaults, so the present value of the protection buyer's payment is

$$U_0^{L,U} + Spr_{L,U}E\left(\sum_{t=1}^{M}(t_i - t_{i-1})DFD(0,t_i)(1 - (L_{i-1}^{L,U} + L_i^{L,U})/2)\right) \tag{3.57}$$

where $t_1, t_2, ..t_M = T_b$. $DFD(0,t_i)$ is the default free discount rate from t_i to time 0. Part of the premium can be paid at time $t_0 = 0$ as an upfront fee $U_0^{L,U}$. According to the market convention, CDX tranches are quoted in terms of upfront fee $U_0^{L,U}$ with a fixed running spread of 100bp; thus, after fixing $Spr_{L,U} = 100bp$, the upfront fee $U_0^{L,U}$ is determined where equation (3.56) = equation (3.57).

Estimation of CDS Index Tranches during 2009

The standard tranche structure of CDX.NA.IG is 0-3% (equity tranche), 3-7% (mezzanine tranche), 7-10% and 10-15%. Table 3.6 shows the summary statistics of daily CDX.NA.IG13 tranche market quotes between September 23, 2009 to December 31, 2009.

We simulate the posterior distribution of each CDX tranche price with the last 100 posterior samples of parameters. The simulation method of CDX tranche price with given (m^{th} sampled)

Table 3.6: **Summary of CDX.NA.IG13 Equity Market Quotes (%) between September 23, 2009 to December 31, 2009 (68 days).**

Tranche	0-3%	3-7%	7-10%	10-15%
Min	48	0	5.813	0.5
1st Qu.	51.7	20.47	7.813	1
Median	55.4	22.12	8.312	1.625
Mean	55.4	21.39	8.215	1.584
3rd Qu.	58.8	23.12	8.922	2.063
Max	62.5	26.62	10.690	3.313

$\ln V_{it}^{(m)}, \sigma_{V_i}^{(m)}$ is as follows: The simulation is carried out by drawing a set of zero mean, unit variance, normally distributed random $\Delta x_{1t}, \Delta x_{2t}, \Delta y_t$ and Δw_{it}. The asset value of firm i at time t_k is then derived as below:

$$\begin{aligned} \ln V_{it_k}^{(m)} &= \ln V_{it}^{(m)} + (r - 1/2\sigma_i^{(m)2})(t_k - t) + \sigma_i^{(m)}(a_{1i}\Delta x_{1t} + a_{2i}\Delta x_{2t} + b_i \Delta x_{yt})\sqrt{t_k - t} \\ &\quad + \sigma_{V_i}^{(m)} \sqrt{1 - a_{1i}^2 - a_{2i}^2 - b_i^2}\Delta w_{it}\sqrt{t_k - t} \end{aligned} \tag{3.58}$$

Firm i is assumed to default at the mid-point of the time interval (t_{k-1}, t_k); if the value of $V_{it_k}^{(m)}$ is below the barrier, then we stop the derivation of V; if not, then the value $V_{it_{k+1}}^{(m)}$ is sampled again. We iterate this derivation 100 times each, for $m = 1900, ...2000$ and for all $i = 1, ...108$, so for each m, 100 sets of 108 firms' $V_{it}^{(m)}$ paths are sampled and we notate them as $(V_{it}^{m,1}, ...V_{it}^{m,100})$. For each set of paths, the number of defaults, that occur in each time interval, (t_{k-1}, t_k), is determined; then the loss $Loss_t$ in equation (3.55) can be estimated by $\hat{L}oss_t$ as below:

$$\hat{Loss}_t^{(m)} = \frac{\sum_{i=1}^{N}(1 - R_i)\frac{1}{100}\sum_{l=1}^{100} 1_{\hat{D}_{it}^{m,l}=1}}{N} \tag{3.59}$$

where $\hat{D}_{it}^{m,l}$ is default indicator of firm i by time t at l^{th} default-time-simulation iteration with m^{th} posterior samples of parameters. We then get the posterior distribution of tranche prices by obtaining $\hat{Loss}_t^{(m)}$ for $m = 1900, ..., 2000$.

Figure 3.19 depicts the structural model results. For comparison, we also get the posterior distribution of the CDO quote when correlation structures are not incorporated and only observed

common factors are incorporated.

Figure 3.18 is the histograms of the posterior distributions of CDO tranche prices on December 31, 2009, under the structural model. As shown in histograms (i) to (l) in Figure 3.18, without consideration of the correlation among asset processes, equity tranche (0-3%) is over-estimated and other tranche values are under-estimated[11]. However, after incorporating common factors, over- or under-estimation problems are resolved. With only observed common factors (depicted in histograms (e) to (h) in Figure 3.18), most of the correlation structures are captured, but the over-estimation problem in equity tranche (0-3%) is not perfectly adjusted. Adding hidden factors (shown in histogram (a) to (d)) improves the pricing ability especially in terms of the tail behaviors in all tranches.

Figure 3.19 is the histogram of posterior distribution of CDO tranche prices under the intensity model. Adding common factors improves the pricing ability, but its effect is marginal.

By fitting the market CDX tranche prices, we can find weaknesses of our correlated intensity model and directions for further improvement. Different from the structural model based on geometric Brownian motion assumption for asset process, in the intensity model, the CIR process is assumed for the underlying default intensity process. Under such a mean reverting type process, default probability, especially during the long term (5 years in this essay) depends more on the mean level than the volatility level. Therefore, our parallel model structure assigning common factors to volatility parts might not be enough to capture the correlation structure of the default intensities.

A model without jump further worsens the problem. The year 2009 was when the financial crisis was at its worst, but it has recovered since. Accordingly, the CDS market suffered a precipitous drop as well as a rise during this period. Under such circumstances, if not incorporating jump lead over-estimated mean (μ_i') and volatility (σ_i') levels of default intensity dynamic. As a result, such an over-estimation of parameters makes first two tranches are more overly valued than in the struc-

[11] Hull and White (2004) showed the sensitivity of each tranche price to the default correlation. For the equity tranche, higher correlation means lower value to someone buying protection. For the mezzanine tranche, the value of the tranche is not particularly sensitive to correlation. For other senior tranches (7% and more), higher correlation means higher value to someone buying protection.

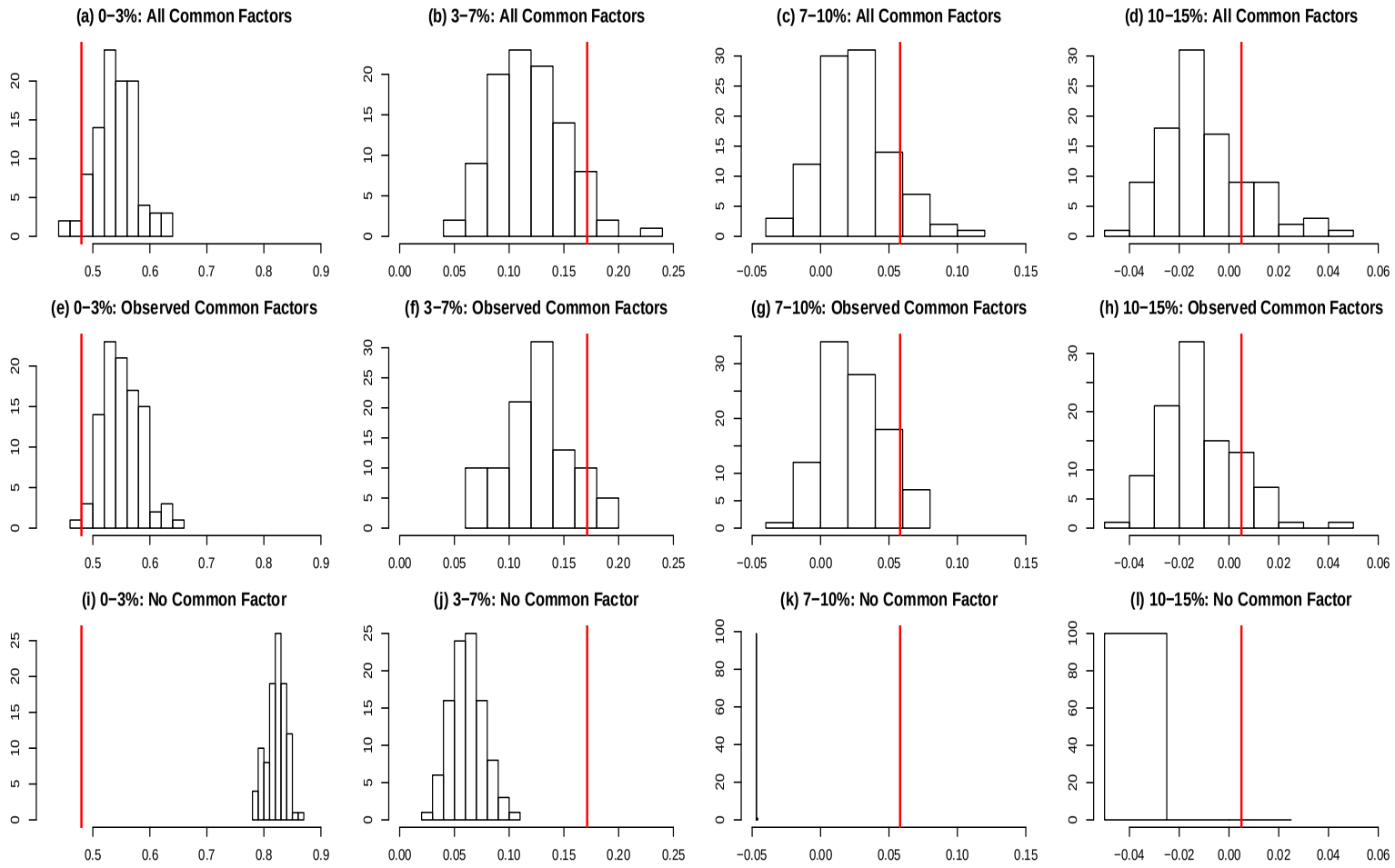

Figure 3.18: **Histograms of the Posterior Samples of CDX Tranche Prices on 31 December, 2009 under the Structural Model.** Histograms (a)-(d) depict the sampled posterior distributions of each tranche price ((from the left) 0-3%, 3-7%, 7-10% and 10-15%), when we consider all common factors. Histograms (e)-(h) depict the sampled posterior distributions of each tranche price, when we consider only observed common factors. Histograms (i)-(l) depict the sampled posterior distributions of each tranche price, when we do not use any of the common factors (independent model).

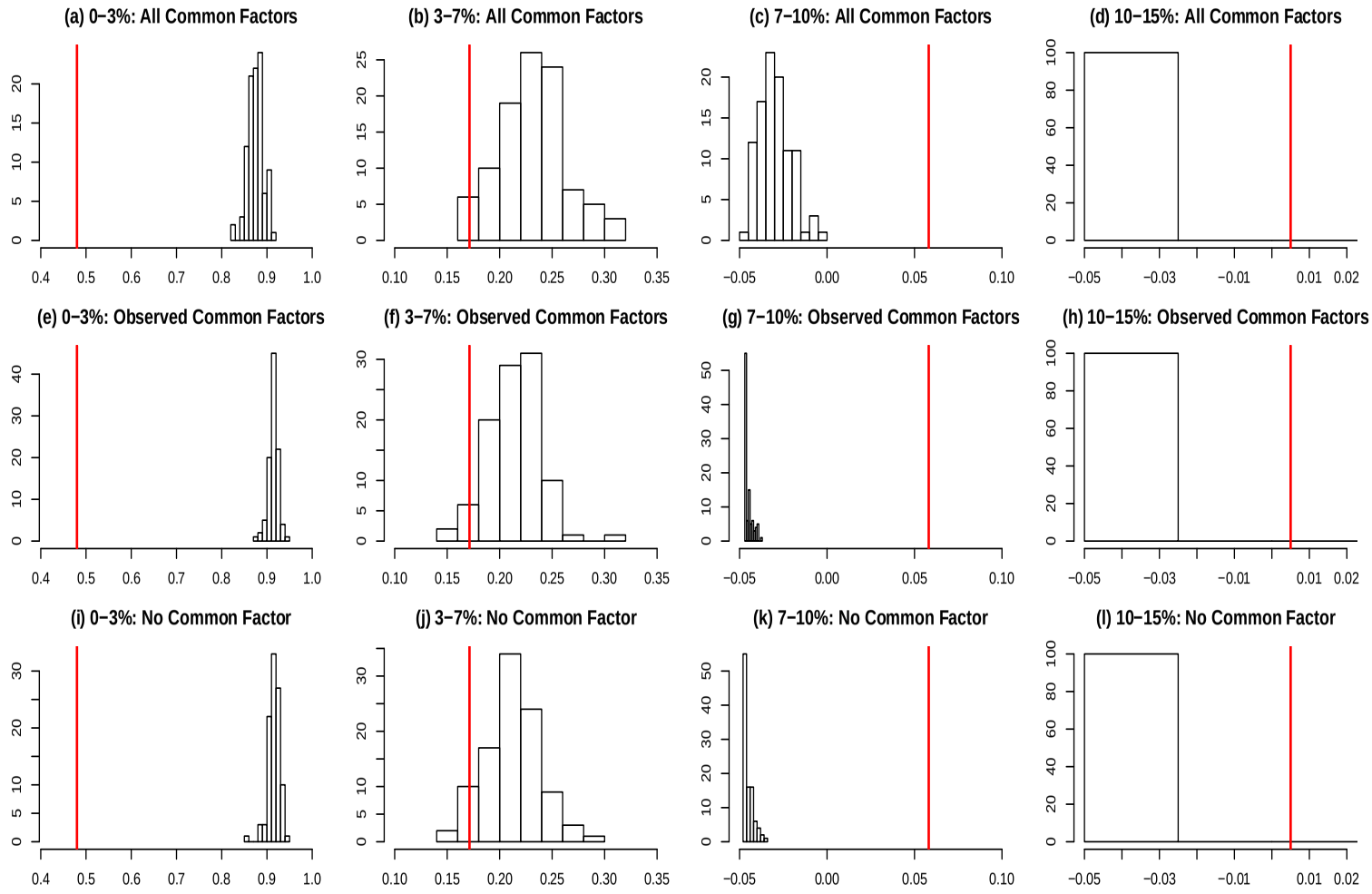

Figure 3.19: **Histograms of the Posterior Samples of CDX Tranche Prices on December 31, 2009 under the Intensity Model.** Histograms (a)-(d) depict the sampled posterior distributions of each tranche price ((from the left) 0-3%, 3-7%, 7-10% and 10-15%), when we consider all common factors. Histograms (e)-(h) depict the sampled posterior distributions of each tranche price, when we consider only observed common factors. Histograms (i)-(l) depict the sampled posterior distributions of each tranche price, when we do not use any of the common factors (independent model).

tural model. Moreover, as noted in Mortensen (2006) and Feldhutter and Nielsen (2012), without jumps, the model is a pure diffusion model and is unable to generate enough default correlation to match senior tranche spreads.

We can improve our correlated intensity model in further research by including correlated jump or time-varying mean level that are dependent on common factors. In terms of the prediction of financial derivative prices, our correlated intensity model did not perform well as much as in our correlated structural model. However, in that we measure the correlated random movement in the intensity movement, our parallel framed model accomplishes significant enhancement.

The CDX tranche price prediction analysis give us another direction of further study. Because our estimation methods are Gibbs Sampling and the MH algorithm, one-step-ahead prediction is hard to achieve. This is why we present in this essay only the result of prediction on December 31. Prediction on all other days under the structural model are presented in appendix B.5. However, the prediction of CDX tranche prices on September 23 with the entire 2009 data set is not practical. However, under the Gibbs and MH methods, when new data comes in, all the estimation process must be done from the beginning. In fact, we attempted prediction of CDX tranche price on September 23 based only on the previous 6-month data range from April 22 to September 22, and likewise on December 31 using the data range from September 30 to December 30. The results (which are given in the appendix B.5) are similar to Figures 3.18 and 3.19. However, in order to predict these two date prices, we need to conduct two independent posterior samplings. Therefore, prediction of daily prices of CDX tranche is not feasible using Gibbs and MH methods. If we can use the Bayesian estimation method, which can make one-step-ahead predictions, then our models can give more useful and interesting insights to financial practitioners.

3.12 Conclusion

In this essay, we proposed a common model frame for both structural and intensity credit risk models by combining the merits of several default correlation studies, which are independently developed under each model framework. By combining both, in the structural model, we improved previous models to be more easily extendable to a large number of firms. In the intensity model,

it became possible to reduce the necessary data size. In order to deal with such combined complicated models, the following Bayesian estimation techniques were applied: generalized Gibbs sampling and MH algorithm.

More than combining model structures, we relaxed several assumptions on which previous studies rely. Through simulation studies and empirical study based on 2009 stock and CDS prices information of 125 firms comprising CDX.NA.IG13, we showed the necessity of following three extensions we proposed in this essay.

First, CDS prices are used in the estimation process. We assumed the state variable, asset or intensity, follows a certain stochastic process and this determines the default time and other financial derivative prices. By doing so, we obtained the model frame, which easily incorporates multiple-securities information in estimation processes and enhances the estimation performance. In fact, both stock and CDS prices information were used together for asset process estimation. The simulation and empirical results show that adding CDS prices in estimation makes it possible to obtain improvements in posterior sample convergence.

Second, we added hidden factors as one of the common factors determining the default correlation. The empirical results show that incorporating hidden factors is necessary in the structural and intensity models because it can explain asset and intensity correlations, which are not captured by the observed common factors. However, their role under each model is different. In the intensity model, the observed macroeconomic factor cannot explain the co-movement of intensities significantly, thus making it necessary to add a hidden factor in order to capture the basic correlation of intensities. In the structural model, even though SP500 and HML both significantly explain the co-movement of asset processes, their explanation is constrained to a certain group of firms, especially those that dominate the S&P500 return during the data period. Incorporating hidden factors therefore, makes it possible to capture more diversified correlation structures through the random coefficients assumption.

Third, we relaxed the assumptions of equal factor effects across entire firms applied in previous studies. The empirical study supports the necessity of this relaxation in assumption and suggests that Duffie and Saita (2009)'s hypothesis that all firms have the same sensitivity is not supported

by the data.

This non-equal common factor effect model suggests one of the potential uses of our model. The non-equal common factor effect makes it possible to derive the pair-wise and firm-by-firm correlation structure. In terms of risk management strategy, the risk of specific portfolios can be more correctly measured. This valuation could not be included in this essay, but instead of CDX tranche, which includes all 125 issuers, there might be a portfolios composed of a small number of bonds. In such cases, inference based on our model can offer better understanding about their specific risk structures, because we derived the firm-by-firm specific correlation structure.

Another potential contribution of the model built in this essay is that more diversified predictions are possible. Since specific stochastic processes on the states are assumed and estimated with consideration of the correlation structure, we can predict the value of states in multiple futures. Moreover, based on the predicted state values, we can predict the price of other financial derivatives, especially those that have basket structures, even though they are not used in the estimation process.

In order to demonstrate such a potential practical application in predictions, the posterior distribution of CDX tranche prices is driven in this essay by taking advantage of Bayesian estimation. Incorporating the correlation structure improved the prediction of CDX tranche market quotes, especially under the correlated structural model. However, in the intensity model, prediction did not work as well because all observed factors only marginally impact the intensity. With one common factor, there might be limitations in fully capturing the correlation structure. However, different from previous CDX tranche price studies, the CDX tranches market information is not incorporated in the process of estimation in this essay. The prediction of CDX tranche prices has also meaning in that it is a model validation check with complete out-of-sample data.

Moreover, the CDX tranches prices prediction result suggests that, with further research, when including jump diffusion in the dynamics of intensity and asset, we can expect improvements in CDX tranche prices predictions. While we combined two types of financial market information for the structural model, for the intensity model, we used only CDS prices. Future research might include bond prices instead of equity prices for the intensity model. As in the structural model,

adding additional information is expected to improve estimation precision and convergence in the intensity model. Moreover, our work in CDX tranche prices prediction necessitate the application of other estimation methods that make a one-step-ahead prediction possible.

APPENDIX

A. APPENDICES TO CHAPTER 1

A.1 Bond Price under the Black-Cox Model

Lando (2004) provides a explicit formula for $B(V(t),t;T_D,D)$,the price of the defaultable bond that pays D if $\tau > T_D$ and pays $L(\tau)$ if $\tau < T_D$.

$B((V(t),t;T_D,D)$ solves the following set of partial differential equations:

$$\frac{\partial B}{\partial t} + rV\frac{\partial B}{\partial V} + \frac{1}{2}\sigma_V^2 B^2 \frac{\partial^2 V}{\partial B^2} - rB = 0 \tag{A.1}$$

$$B(L(t),t;T_D,D) = L(t) \tag{A.2}$$

$$B(V,T_D;T_D,D) = \min(D,V) \tag{A.3}$$

$$B(\infty,t;T_D,D) = D\exp(-r(T_D-t)). \tag{A.4}$$

These partial differential equations can be solved numerically. As specified in Black and Cox (1976), $B((V(t),t;T_D,D)$ can be obtained by applying the risk-neutral valuation principle:

$$B(V(t),t;T_D,D) = E^Q\left[e^{-r\tau}L(\tau)\mathbf{1}_{\{\tau\le T_D\}}\right] + E^Q\left[e^{-r(T_D-t)}\min(V(T_D),D)\mathbf{1}_{\{\tau>T_D\}}\right]. \tag{A.5}$$

In order to compute the risk-neutral valuation formula above, we need to obtain the joint density of a Brownian motion with a constant drift and its running minimum. For some constant α, we define a Brownian motion $\hat{W}(t)$ with a drift α and its running minimum by:

$$\hat{W}(t) = \alpha t + W_V^Q(t),\ \ m(t) = \min_{0\le s\le t}\hat{W}(s). \tag{A.6}$$

By applying the reflection principle of Brownian motion and changing the measure to remove the drift term, one can derive $f_{\hat{W}(t),m(t)}(w,m)$, the joint density of $(\hat{W}(t),m(t))$, as follows:

$$f_{\hat{W}(t),m(t)}(w,m) = \frac{2(w-2m)}{t\sqrt{2\pi t}} e^{\alpha w - \frac{1}{2}\alpha^2 t - \frac{(w-2m)^2}{2t}}, \;\; w \geq m, \;\; m \leq 0. \tag{A.7}$$

Lando (2004) carried out expectation calculations with the density given above. The first expectation in equation (A.5) is

$$E^Q\left[e^{-r\tau}L(\tau)\mathbf{1}_{\{\tau \leq T_D\}}\right] \;\; = \;\; L(0)\frac{\exp(b\mu)}{\exp(b\tilde{\mu})}\left[N\left(\frac{b-\tilde{\mu}(T_D-t)}{\sqrt{T_D-t}}\right) + \exp(2\tilde{\mu}b)N\left(\frac{b+\tilde{\mu}(T_D-t)}{\sqrt{T_D-t}}\right)\right], \tag{A.8}$$

where

$$b = \frac{\left(\frac{L(0)}{V(t)}\right) - \gamma(T_D-t)}{\sigma_V}, \mu = \frac{r - \frac{1}{2}\sigma_V^2 - \gamma}{\sigma_V}, \;\; \text{and } \tilde{\mu} = \sqrt{\mu^2 + 2r}.$$

The second expectation in equation (A.5) is

$$\begin{aligned}
&E^Q\left[e^{-r(T_D-t)}\min(V(T_D),D)\mathbf{1}_{\{\tau > T_D\}}\right] \qquad (A.9)\\
&= L(0)\exp(-\gamma(T_D-t))\exp(-r(T_D-t))N\left(\frac{\left(\frac{V(t)}{L(0)}\right) + \left(r - \frac{1}{2}\sigma_V^2\right)(T_D-t)}{\sigma_V\sqrt{T_D-t}}\right)\\
&\quad -L(0)\exp(-\gamma(T_D-t))\left(\frac{L(0)}{V(t)}\right)^{\frac{2r}{\sigma_V^2}-1}\exp(-r(T_D-t))N\left(\frac{-\ln\left(\frac{V(t)}{L(0)}\right) + \left(r - \frac{1}{2}\sigma_V^2\right)(T_D-t)}{\sigma_V\sqrt{T_D-t}}\right)\\
&\quad +C^{BS}(\exp(-\gamma(T_D-t))V(t),t;T_D,L(0)\exp(-\gamma(T_D-t)))\\
&\quad -\left(\frac{L(0)}{V(t)}\right)^{\frac{2r}{\sigma_V^2}-1}C^{BS}\left(\frac{L(0)^2\exp(-\gamma(T_D-t))}{V(t)},t;T_D,L(0)\exp(-\gamma(T_D-t))\right)\\
&\quad -C^{BS}(\exp(-\gamma(T_D-t))V(t),t;T_D,D\exp(-\gamma(T_D-t)))\\
&\quad +\left(\frac{L(0)}{V(t)}\right)^{\frac{2r}{\sigma_V^2}-1}C^{BS}\left(\frac{L(0)^2\exp(-\gamma(T_D-t))}{V(t)},t;T_D,D\exp(-\gamma(T_D-t))\right).
\end{aligned}$$

A.2 Particle Filter Algorithm

Following S.L. (2001), with known model parameters, we face the non-linear filtering problem of computing $\pi_t(x_t) = P(x_t|y_{1:t})$, the conditional distribution of the state variable X_t given the observed variables $y_{1:t} = (y_1, y_2, \ldots, y_t)$. The posterior distribution of x_t at time t ($\pi_t(x_t)$) can be

recursively derived as,

$$\pi_t(x_t) = P(x_t|y_1,...y_t) \quad (A.10)$$

$$= \int q_t(x_t|x_{t-1}) f_t(y_t|x_t) \pi_{t-1}(x_{t-1}) dx_{t-1} \quad (A.11)$$

where $\pi_{t-1} = p(x_{t-1}|y_1,...y_{t-1})$ is the posterior distribution of x_{t-1} at time $t-1$.

In order to generate a set of particles distributed according to $\pi_{t+1}(x_{t+1})$, we adopt the particle filtering method proposed by Gordon, Salmond, and Smith (1993), which makes use of the idea of sampling/importance re-sampling (SIR) introduced by Rubin (1987). The algorithm is given below:

1. Suppose that at time t, we have a random sample $\{x_t^{(1)},....,x_t^{(m)}\}$, which follow approximately $\pi_t(x_t)$
2. Draw $x_{t+1}^{(*j)}$ from the state equation $q_{t+1}(x_{t+1}|x_t^{(j)}), j = 1,...,m$
3. Weight each draw by $w^j \approx f_{t+1}(y_{t+1}|x_{t+1}^{*j})$
4. Resample from $\{x_{t+1}^{(*1)},....,x_{t+1}^{(*m)}\}$ with probability proportional to $w^{(j)}$ to produce a random sample $\{x_{t+1}^{(1)},....,x_{t+1}^{(m)}\}$ for time $t+1$

B. APPENDICES TO CHAPTER 3

B.1 The Posterior Scale Reduction Factor

B.1.1 The Convergence of Model 1

Table B.1: **The Posterior Scale Reduction Factor (PSRF) of Common Factor Coefficients and Asset Volatilities in the Simulated Structural Model.** These are the PSRF of posterior samples of common factor coefficients and asset volatilities, when we use both CDS and stocks to estimated the simulated structural model.

	Firm 1	Firm 2	Firm 3	Firm 4	Firm 5	Firm 6	Firm 7	Firm 8	Firm 9	Firm 10
$a1_i$	1.0083	1.0006	1.0056	1.0009	1.0031	1.0015	1.0003	1.0057	1.0017	1.0011
b_i	1.0021	0.9996	1.0011	0.9999	0.9995	0.9998	0.9999	1.0118	0.9996	1.0024
σ_i	0.9995	0.9995	0.9995	0.9995	0.9996	0.9995	0.9995	0.9995	0.9995	0.9995

Table B.2: **Summary Statistics of the Posterior Scale Reduction Factor of Hidden Factor Process (Length =252 days) in the Simulated Structural Model.** These are summary statistics of the PSRF of posterior samples of hidden factor process (one year daily values, length=252), when we use both CDS and stocks to estimate the simulated structural model.

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.9995	1.0130	1.0320	1.0460	1.0680	1.2580

Table B.3: **Summary Statistics of the Posterior Scale Reduction Factor of Asset Value Processes (10 firms $\times$ 252 days) in the Simulated Structural Model.** These are summary statistics of the PSRF of the posterior samples of asset value processes (10 firms' one year daily values, length=2520), when we use both CDS and stocks to estimate the simulated structural model.

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.9995	1.0010	1.0040	1.0130	1.0090	1.3520

B.1.2 The Convergence of Model 2

Table B.4: **The Posterior Scale Reduction Factor (PSRF) of Common Factor Coefficients and Asset Volatilities in the Simulated Intensity Model.** These are the PSRF of posterior samples of common factor coefficients and intensity volatilities.

	Firm 1	Firm 2	Firm 3	Firm 4	Firm 5	Firm 6	Firm 7	Firm 8	Firm 9	Firm 10
$a1_i$	1.0202	1.0001	1.0061	1.0046	1.0000	1.0007	1.0012	1.0263	1.0108	1.0096
b_i	1.0050	1.0002	1.0020	0.9996	0.9996	1.0006	1.0001	1.0017	1.0161	1.0015
σ_i	0.9998	0.9995	1.0016	0.9997	0.9995	1.0003	1.0007	1.0061	1.0055	1.0108

Table B.5: **Summary Statistics of the Posterior Scale Reduction Factor of Hidden Factor Process (Length =252 days) in the Simulated Intensity Model.** These are summary statistics of the PSRF of posterior samples of hidden factor process (one year daily values, length=252) in the simulated intensity model.

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1.000	1.040	1.089	1.129	1.174	2.028

Table B.6: **Summary Statistics of the Posterior Scale Reduction Factor of Default Intensity Processes (10 firms $\times$ 252 days)in the Simulated Intensity Model.** These are summary statistics of the PSRF of posterior samples of default intensity processes (10 firms' one year daily values, length=2520).

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.9996	1.0370	1.0740	1.1030	1.1390	1.7530

B.2 More Simulation Results

B.2.1 More Simulation Results for Model 1

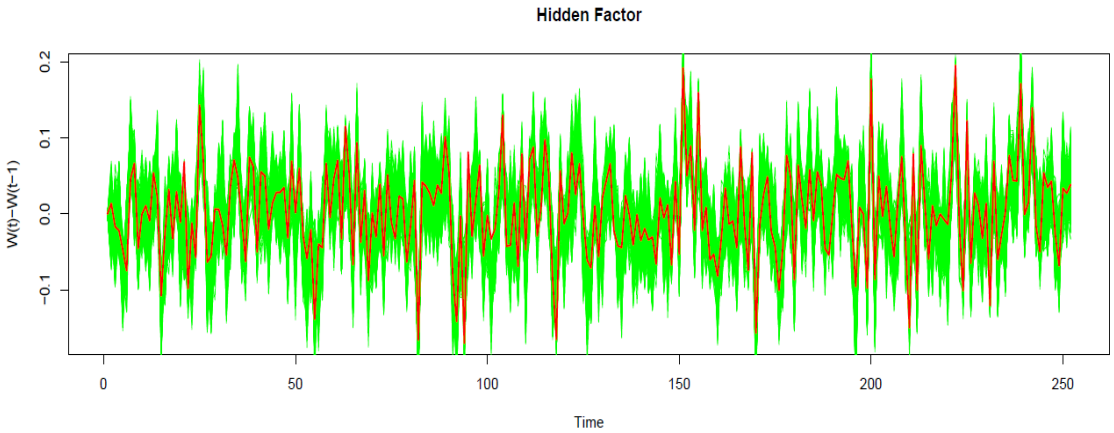

Figure B.1: **Posterior Samples of Hidden Factor Values $(W_t - W_{t-1})$ Using Only Stock Prices in the Simulated Structural Model.**

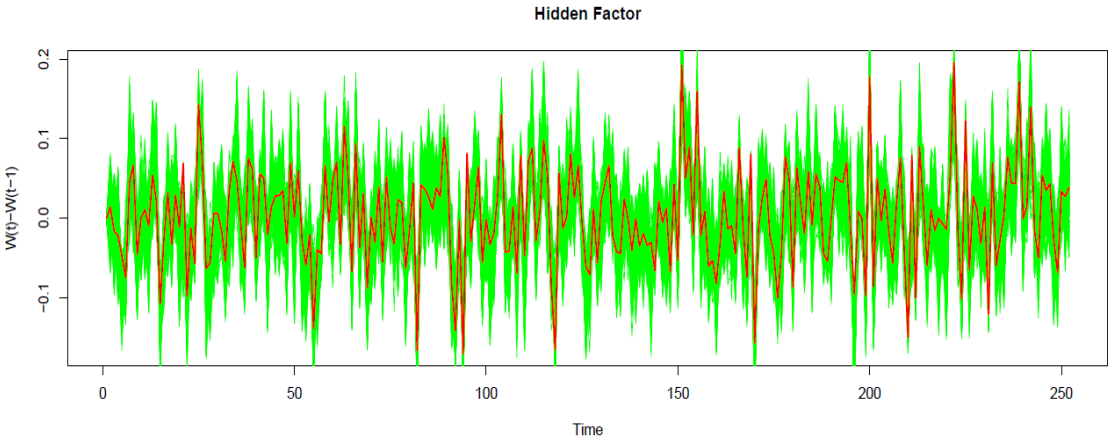

Figure B.2: **Posterior Samples of Hidden Factor Values $(W_t - W_{t-1})$ Using Both Stock and CDS Prices in the Simulated Structural Model.**

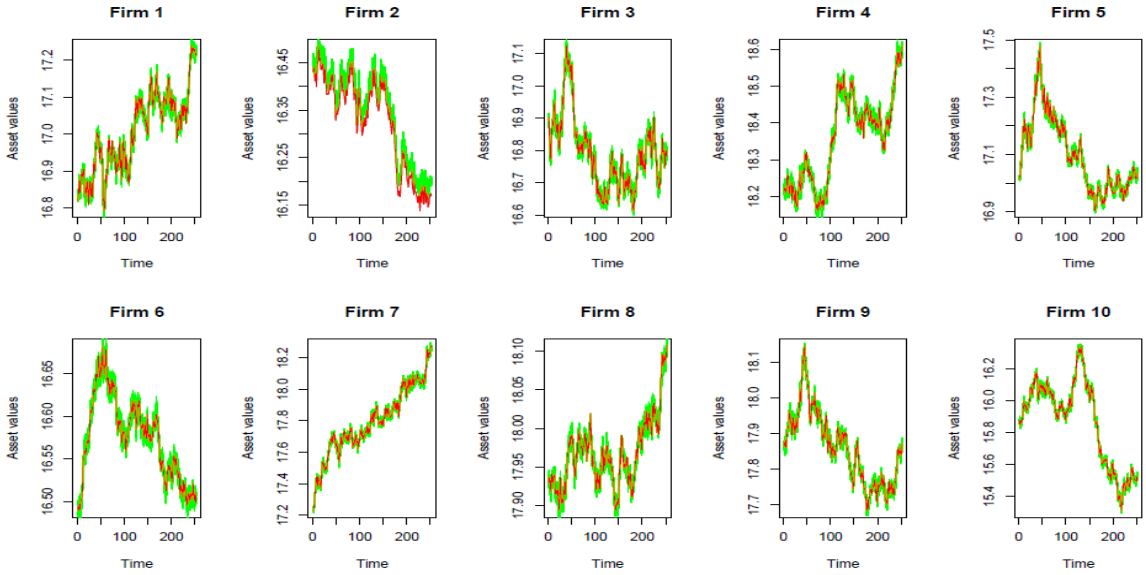

Figure B.3: **Posterior Samples of Asset Processes $\ln V_{it}$ for Firms 1-10 When We Use Only Stock Prices in the Simulated Structural Model.**

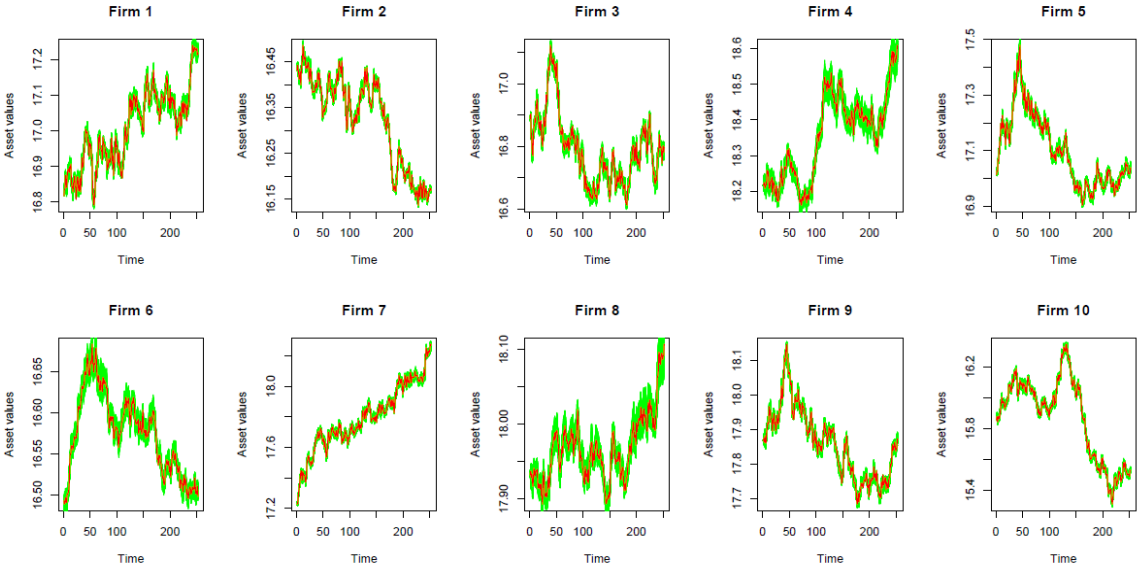

Figure B.4: **Posterior Samples of Asset Processes $\ln V_{it}$ for Firm 1-10 When We Use Both Stock and CDS Prices in the Simulated Structural Model.**

B.2.2 More Simulation Results for Model 2

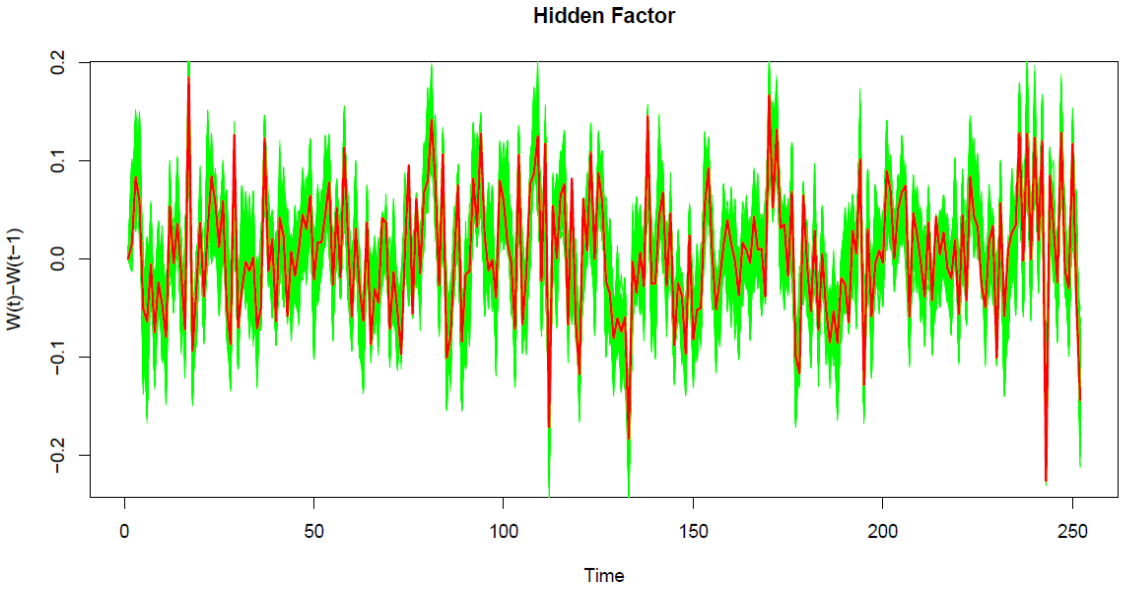

Figure B.5: **Posterior Samples of Hidden Factor Values $(W_t - W_{t-1})$ in the Simulated Intensity Model.**

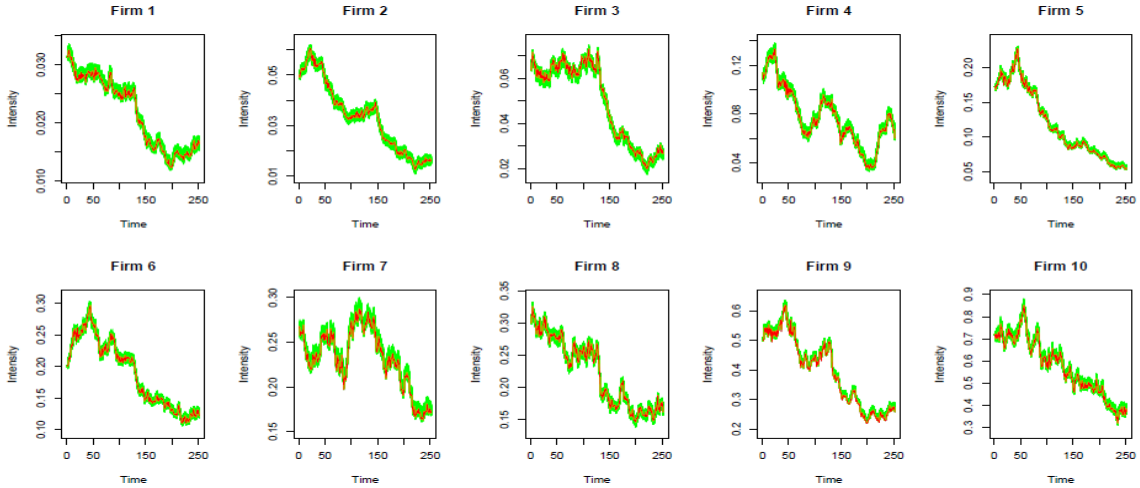

Figure B.6: **Posterior Samples of Intensity Processes λ_{it} for Firms 1-10 in the Simulated Intensity Model.**

B.3 Summary of 108 Companies in the Sample

Table B.7: **Summaries of 108 Companies Included**

COMPANY	TICKER	MEAN.CDS	SD.CDS	MEAN.S	SD.S
ACE Ltd.	ACE	99.14	21.12	46.71	5.04
Aetna Inc.	AET	94.34	26.37	27.33	3.09
Alcoa Inc.	AA	462.64	246.58	10.82	2.7
Allstate Corp./The	ALL	155.55	97.02	25.86	4.25
Altria Group Inc.	MO	102.31	21.71	17.37	1.18
American Electric Power Co., Inc.	AEP	65.39	11.84	29.85	2.81
American Express Co.	AXP	259.46	161.43	26.97	9.19
American International Group Inc.	AIG	1220.33	687.77	16.74	17.01
Amgen Inc.	AMGN	63.04	15.79	54.91	4.68
Anadarko Petroleum Corp.	APC	146.09	83.71	49.82	10.13
Arrow Electronics Inc.	ARW	103.71	41.65	23.43	4.02
AT& T Inc.	T	104.65	26.65	25.62	1.29
AutoZone Inc.	AZO	80.35	27.72	150.39	9.14
Avnet Inc.	AVT	190.47	111.35	22.84	3.86
Barrick Gold Corp.	ABX	131.45	72.65	35.59	4.14
Baxter International Inc.	BAX	32.4	8.14	54.02	3.42
Black & Decker Corp.	BDK	151.13	65.93	40.31	12.42
Bristol-Myers Squibb Co.	BMY	40.06	11.97	21.72	1.83
Campbell Soup Co.	CPB	31.15	7.73	30.16	2.67
Cardinal Health Inc.	CAH	55.9	12.35	32.24	3.21
Carnival Corp.	CCL	202.72	101.73	27.08	4.68
Caterpillar Inc.	CAT	173.32	87.24	42.02	11.09
CBS Corp.	CBS	286.76	130.36	8.83	3.27
CenturyLink Inc.	CTL	90.65	20.08	30.65	3.24
Chubb Corp.	CB	64.85	23.38	44.58	4.64
CIGNA Corp.	CI	177.3	38.79	24.64	6.11
Cisco Systems Inc.	CSCO	63.54	27.3	19.99	3.13
Comcast Corp.	CMCSA	146.95	34.75	14.92	1.3
Computer Sciences Corp.	CSC	56.8	21.8	44.95	7.35
ConAgra Foods Inc	CAG	45.92	13.01	19.23	2.36
ConocoPhillips	COP	64.08	24.62	45.5	4.81
Constellation Energy Group Inc.	CEG	258.86	95.97	27.78	4.75
CSX Corp.	CSX	89.42	42.55	37.03	8.44
CVS Caremark Corp.	CVS	65.25	20.13	31.44	3.41
Darden Restaurants Inc.	DRI	157.22	54.6	32.66	3.42
Deere & Co.	DE	97.87	38.48	42.12	6.78
Dell Inc.	DELL	120.52	75.02	12.62	2.36
Devon Energy Corp.	DVN	69.76	19.92	60.15	8.43
Dominion Resources Inc./VA	D	56.81	11.04	33.6	2.65
Dow Chemical Co./The	DOW	292.91	174.45	18.48	6.75
Duke Energy Corp.	DUK	51.43	8.13	15.07	1.17
Eastman Chemical Co.	EMN	93.85	46.75	42.7	12.86
EI du Pont de Nemours & Co.	DD	87	37.82	28.35	4.71
FirstEnergy Corp.	FE	129.71	27.56	43.29	4.12
Fortune Brands Inc.	FO	179.83	46.51	36.6	5.94
GATX Corp.	GMT	263.92	127.56	25.93	3.65
General Mills Inc.	GIS	55.25	20.1	58.7	6.25

Continued on next page

Table B.7 – continued from previous page

COMPANY	TICKER	MEAN.CDS	SD.CDS	MEAN.S	SD.S
Goodrich Corp.	GR	54.2	15.91	49.27	8.85
Halliburton Co.	HAL	55.32	13.42	22.96	4.84
Hartford Financial Services Group Inc.	HIG	504.62	253.75	17.02	6.86
Hewlett-Packard Co.	HPQ	46.37	18.75	40.4	6.92
Home Depot Inc.	HD	114.3	58.91	25.01	2.52
Honeywell International Inc.	HON	62.04	21.3	34.2	3.99
Ingersoll-Rand Co.	IR	81.04	23.7	24.84	7.69
International Business Machines Corp.	IBM	54.53	22.45	109.28	13.92
International Paper Co.	IP	332.55	218.49	16.4	6.92
Johnson Controls Inc.	JCI	358.03	260.85	20.98	5.67
Kinder Morgan Energy Partners LP	KMP	162.69	55.56	51.54	4.22
Kohl's Corp.	KSS	131.46	67.09	47.33	7.46
Kraft Foods Inc.	KFT	77.32	22.3	26.02	2.03
Kroger Co./The	KR	90.33	12.23	21.95	1.29
Lockheed Martin Corp.	LMT	40.98	9.91	76.26	5.41
Loews Corp.	L	61.89	12.54	29.04	5.27
Lowe's Cos Inc.	LOW	91.26	34.24	20.23	2.12
Marriott International Inc./DE	MAR	257.1	155	22.11	4.26
Marsh & McLennan Cos Inc.	MMC	59.74	12.99	21.46	2.02
McDonald's Corp.	MCD	40.09	13.96	57.44	3.09
McKesson Corp.	MCK	40.17	12.57	48.79	9.92
MDC Holdings Inc.	MDC	105.25	22.08	32.15	3.23
MetLife Inc.	MET	450.1	208.47	31.33	5.83
Newell Rubbermaid Inc.	NWL	179.85	49.63	11.4	3.25
Nordstrom Inc.	JWN	273.33	177.03	24.18	7.94
Norfolk Southern Corp.	NSC	62.34	21.82	42	6.81
Northrop Grumman Corp.	NOC	51.09	10.11	48.14	4.64
Omnicom Group Inc.	OMC	141.24	102.09	32.06	4.82
Pfizer Inc.	PFE	68.37	26.56	15.73	1.7
Progress Energy Inc.	PGN	59.42	8.39	37.62	2.21
Quest Diagnostics Inc./DE	DGX	69.71	23.61	52.97	4.12
Raytheon Co.	RTN	50.9	10.14	46.09	4.28
Reynolds American Inc.	RAI	251.44	75.04	42.56	5.47
RR Donnelley & Sons Co.	RRD	354.55	139.76	14.76	5.17
Ryder System Inc.	R	212.82	89.64	33.87	7.08
Safeway Inc.	SWY	83.12	11.95	20.66	1.49
Sara Lee Corp.	SLE	47.96	14.41	9.95	1.41
Sempra Energy	SRE	89.06	16.82	48.25	4.31
Sherwin-Williams Co./The	SHW	68.77	42.1	55.95	5.24
Southwest Airlines Co.	LUV	228.58	72.72	7.96	1.48
Staples Inc.	SPLS	156.58	76.11	20.52	2.64
Target Corp.	TGT	98.36	39.79	41.05	6.71
Time Warner Cable Inc.	TWC	187.32	69.76	32.1	8.77
Time Warner Inc.	TWX	79.01	28.95	22.92	8.35
TJX Cos Inc.	TJX	77.89	41.63	31.26	6.4
Toll Brothers Inc.	TOL	150.17	29.1	18.9	1.88
Transocean Inc.	RIG	128.48	75	73.3	11.86
Union Pacific Corp.	UNP	66.59	22.39	53.24	8.75
United Parcel Service Inc.	UPS	56.69	22.74	52.17	4.95
UnitedHealth Group Inc.	UNH	208.36	54.58	26.22	3.04
Universal Health Services Inc.	UHS	101.64	54.56	49.63	10.24
Valero Energy Corp.	VLO	238.74	38.37	19.21	2.39

Continued on next page

Table B.7 – continued from previous page

COMPANY	TICKER	MEAN.CDS	SD.CDS	MEAN.S	SD.S
Verizon Communications Inc.	VZ	78.33	24.62	30.46	1.4
Viacom Inc.	VIA	167.63	92.19	22.86	5.17
Wal-Mart Stores Inc.	WMT	66.85	23.26	50.55	2.07
Walt Disney Co./The	DIS	59.52	17.07	24.5	4.43
Wells Fargo & Co.	WFC	142.56	61.09	23.43	5.41
Whirlpool Corp.	WHR	274.86	128.31	52.11	17.73
Xerox Corp.	XRX	276.24	112.88	7.15	1.28
XL Group Plc	XL	464.66	294.14	11.6	5.54
XTO Energy Inc.	XTO	146.27	69.82	39.05	4.36

B.4 More Empirical Results

B.4.1 More Empirical Results for Model 1.

Figure B.7 shows samples of the posterior samples of the hyperparameters and their convergence in Model 1. Hyperparameters of hidden factor effects μ_b and σ_b do not converge as well as hyperparameters of observed common effects coefficients. This is because most of b_i values are near their truncated point 0 and they are widely spread. However, the conditional mean and conditional standard deviation of these truncated normal hyperpriors are all converge well.(Figure B.8). Conditional mean and standard deviation of truncated normal distribution are given in equation(3.53) and (3.54).

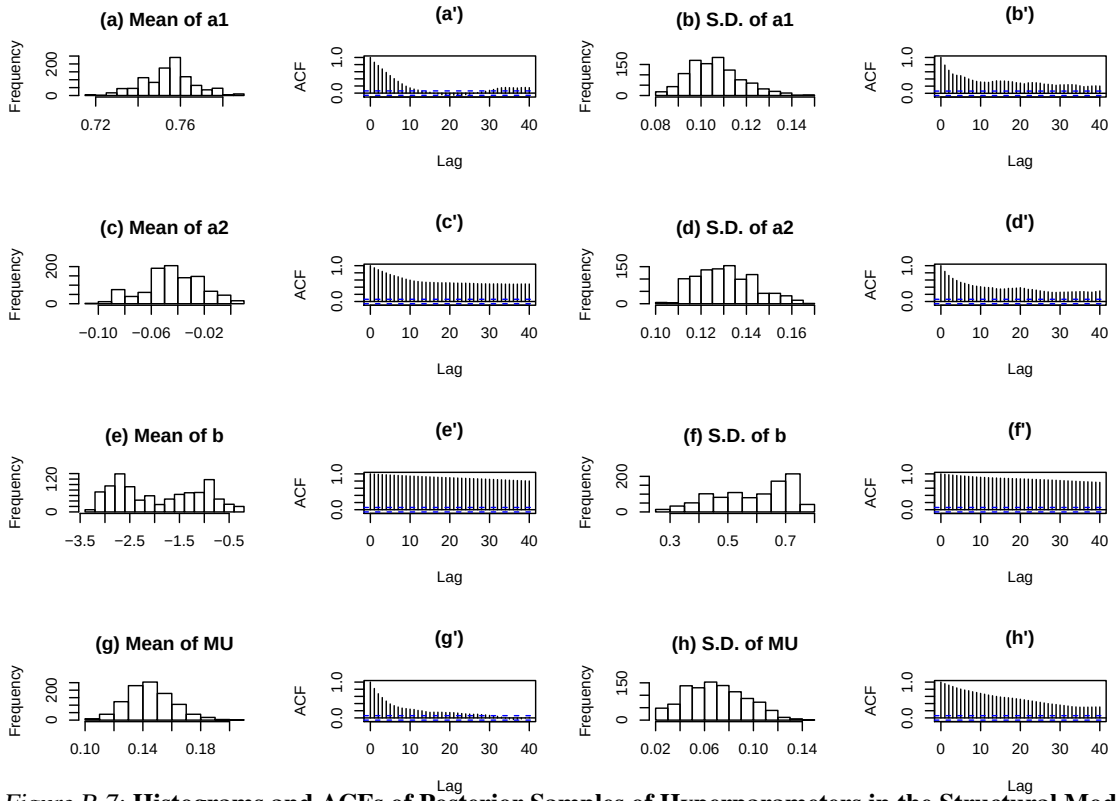

Figure B.7: **Histograms and ACFs of Posterior Samples of Hyperparameters in the Structural Model When We Use Both CDS and Stock Price Information.** Here are histograms and ACFs of posterior samples of hyperparameters. The top row of graphs are for hyperparameters of SP500 coefficients: (a) and (a)' are for the mean (μ_{a1}) and (b) and (b)' are for the standard deviation (σ_{a1}). The second row of graphs are for hyperparameters of HML coefficients: (c) and (c)' are for the mean (μ_{a2}) and (d) and (d)' are for the standard deviation (σ_{a2}). The third row of graphs are for hyperparameters of hidden factor coefficients: (e) and (e)' are for the mean (μ_b) and (f) and (f)' are for the standard deviation (σ_b). The fourth row of graphs are for hyperparameters of mean levels of asset processes: (g) and (g)' are for the mean (μ_μ) and (h) and (h)' are for the standard deviation (σ_μ.)

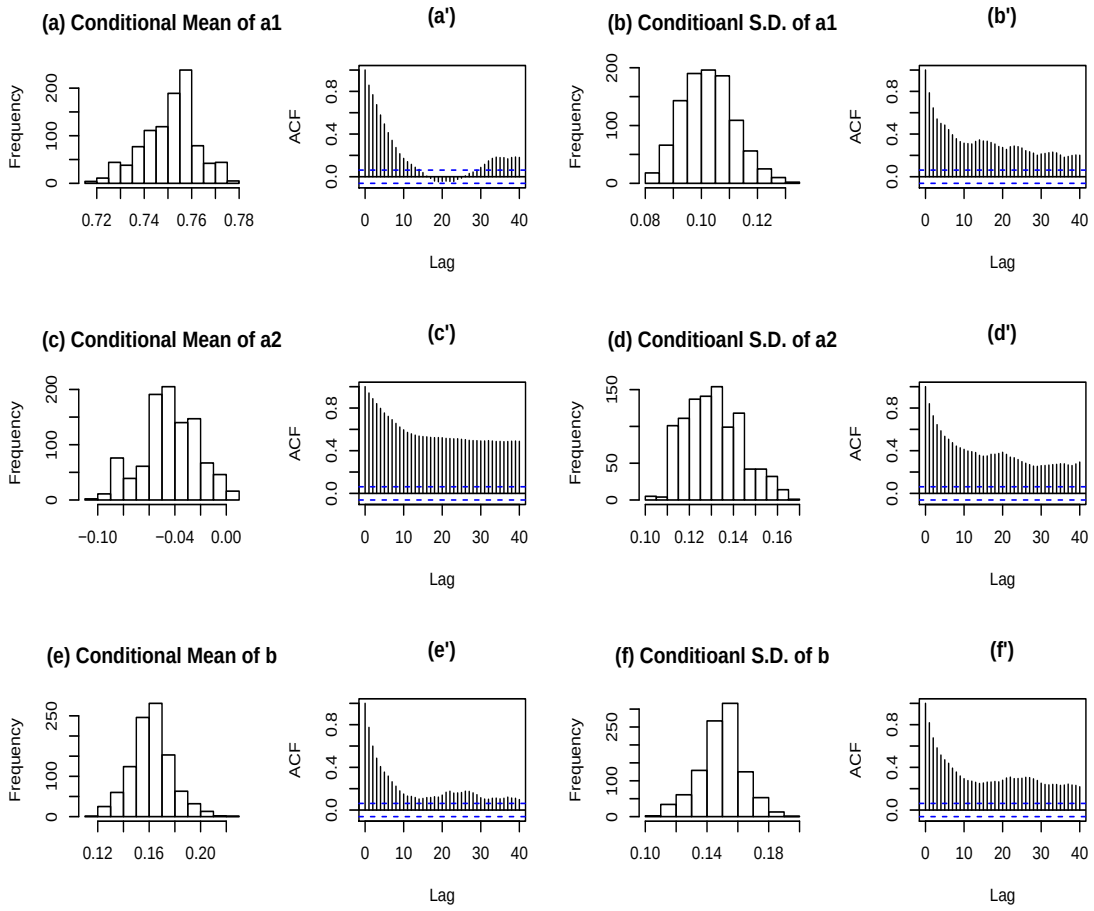

Figure B.8: **Histograms and ACFs of Posterior Samples of Conditional Mean and Standard Deviation of Truncated Normal Prior Distributions in the Structural Model When We Use Both CDS and Stock Price Information.** Here we calculate the conditional mean and standard deviation of truncated normal priors from the posterior samples of hyperparameters. Conditional mean and standard deviation of truncated normal distribution are given in equation(3.53) and (3.54). The top row of graphs are for conditional mean and standard deviation of SP500 coefficients: (a) and (a)' are for the conditional mean and (b) and (b)' are for the conditional standard deviation. The second row of graphs are for conditional mean and standard deviation of HML coefficients: (c) and (c)' are for the conditional mean and (d) and (d)' are for the conditional standard deviation. The third row of graphs are for conditional mean and standard deviation of hidden factor effect coefficients: (e) and (e)' are for the conditional mean and (f) and (f)' are for the conditional standard deviation.

With the posterior means of coefficients of SP500, HML, and the hidden factor, a 108×108 correlation matrix is derived by the equation (3.25). With posterior mean of each common factor coefficients a_{1i}, a_{2i} and b_i, we calculate correlation coefficient of pair of firms. Figure B.9 is a boxplot of each row in the correlation matrix.

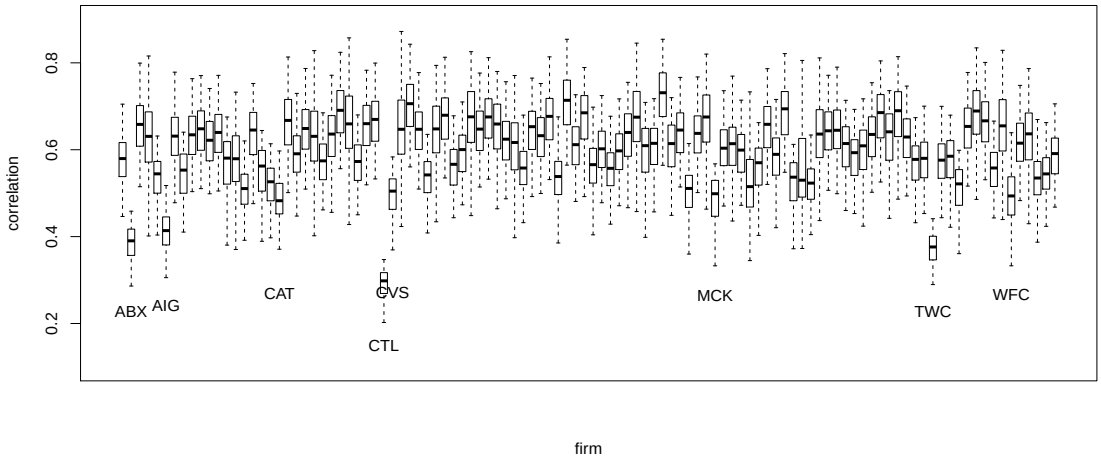

Figure B.9: **Estimates of Correlation Coefficients in the Structural Model.** This is a boxplot of each row in the correlation matrix derived by the equation (3.25). We apply posterior mean of a_{1i}, a_{2i} and b_i for the correlation coefficient calculation.

B.4.2 More Empirical Results for Model 2.

Figure B.10 shows estimates of posterior mean and convergence of posterior sample of hyperparameters in Model 2. The conditional mean and conditional standard deviation of truncated normal hyperpriors are all converge well.(Figure B.11).

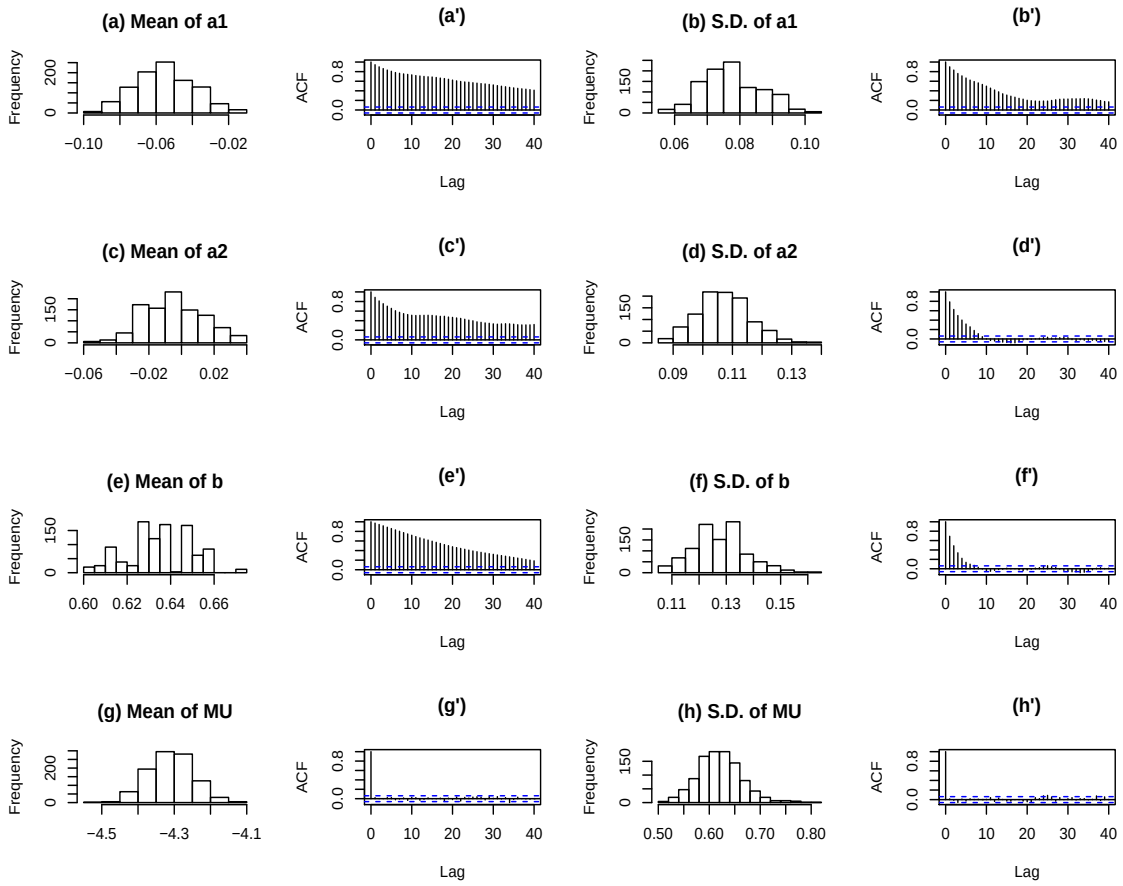

Figure B.10: **Histograms and ACFs of Posterior Samples of Hyperparameters in the Intensity Model.** Here are histograms and ACFs of posterior samples of hyperparameters. The top row of graphs are for hyperparameters of SP500 coefficients: (a) and (a)' are for the mean (μ_{a1}) and (b) and (b)' are for the standard deviation (σ_{a1}). The second row of graphs are for hyperparameters of HML coefficients: (c) and (c)' are for the mean (μ_{a2}) and (d) and (d)' are for the standard deviation (σ_{a2}). The third row of graphs are for hyperparameters of hidden factor coefficients: (e) and (e)' are for the mean (μ_b) and (f) and (f)' are for the standard deviation (σ_b). The fourth row of graphs are for hyperparameters of log-mean levels of intensity processes: (g) and (g)' are for the mean (μ_μ) and (h) and (h)' are for the standard deviation (σ_μ.)

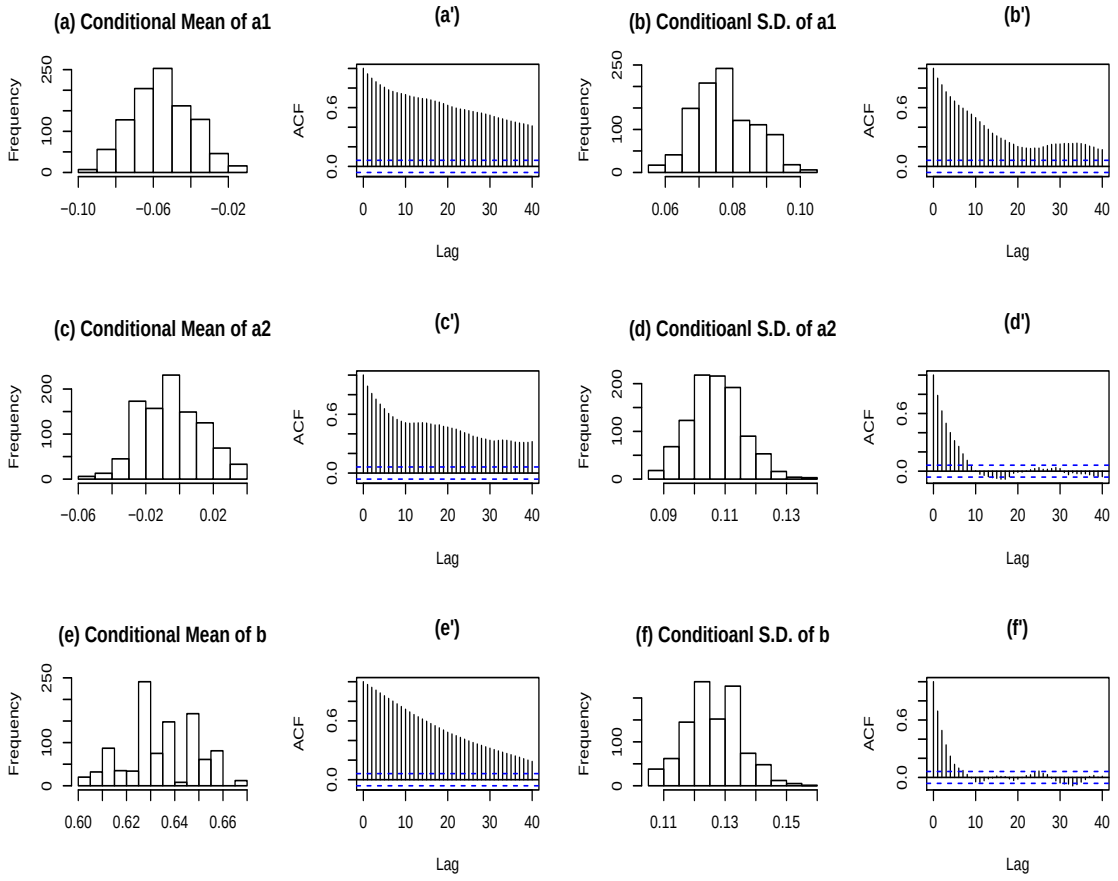

Figure B.11: **Histograms and ACFs of Posterior Samples of Conditional Mean and Standard Deviation of Truncated Normal Prior Distributions in the Intensity Model.** Here we calculate the conditional mean and standard deviation of truncated normal priors from the posterior samples of hyperparameters. Conditional mean and standard deviation of truncated normal distribution are given in equation(3.53) and (3.54). The top row of graphs are for conditional mean and standard deviation of SP500 coefficients: (a) and (a)' are for the conditional mean and (b) and (b)' are for the conditional standard deviation. The second row of graphs are for conditional mean and standard deviation of HML coefficients: (c) and (c)' are for the conditional mean and (d) and (d)' are for the conditional standard deviation. The third row of graphs are for conditional mean and standard deviation of hidden factor effect coefficients: (e) and (e)' are for the conditional mean and (f) and (f)' are for the conditional standard deviation.

The convergence is not as good as in the structural model because, as mentioned before, the effects of observed factors are marginal compared to the hidden factor effect. Hyperparameters governing the distribution of the hidden factor effect and the log-transformed mean level of intensity process all converge well.

Figure B.12 is a boxplot of each row in an estimated intensity correlation matrix. AIG has the smallest intensity correlation with others as in the structural model. However, in general, the size of intensity correlation coefficients are smaller than for the asset correlation.

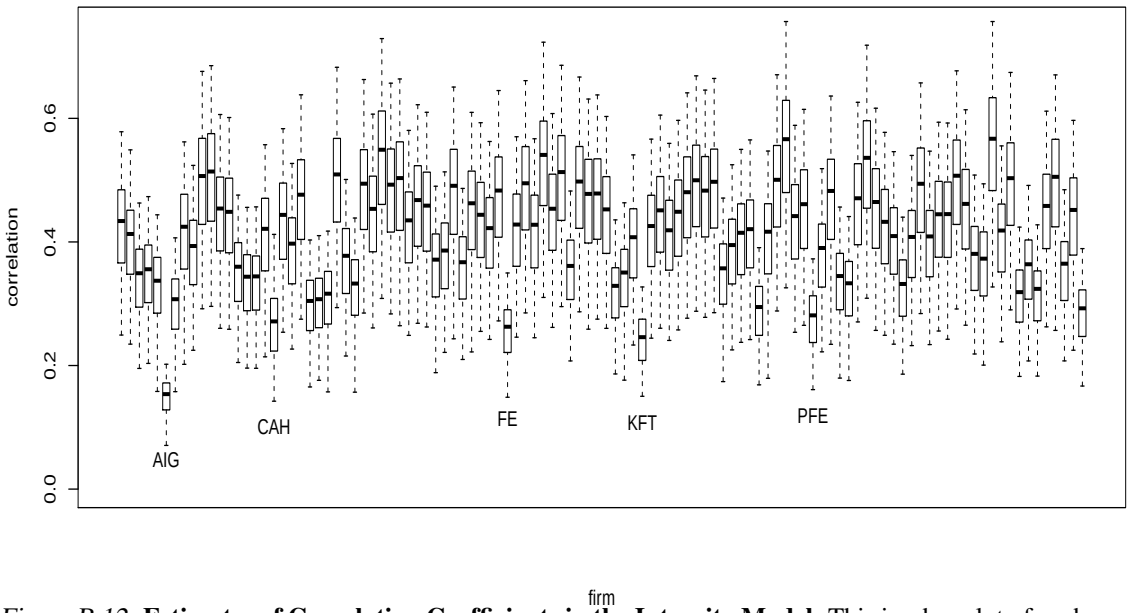

Figure B.12: **Estimates of Correlation Coefficients in the Intensity Model.** This is a boxplot of each row in the correlation matrix derived by the equation (3.34). We apply posterior mean of a_{1i}, a_{2i} and b_i for the correlation coefficient calculation.

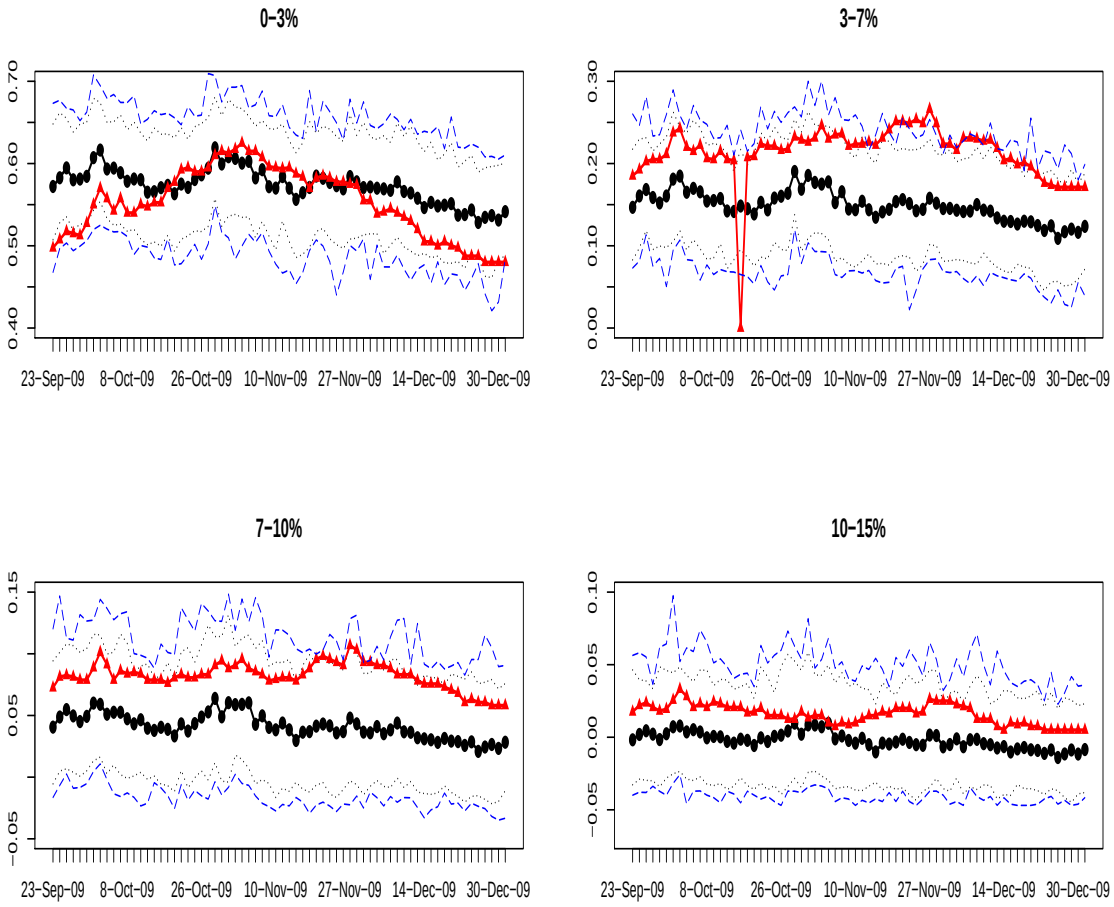

Figure B.13: **Sampled CDX Tranche Prices under the Structural Model.** We derive the posterior predictive distribution for all the daily CDX tranche quotes based on estimated results described in section 3.10. Black lines are the posterior mean paths. Red lines are market quotes. Blue dotted lines are 95% posterior intervals and blue dashed lines show the ranges.

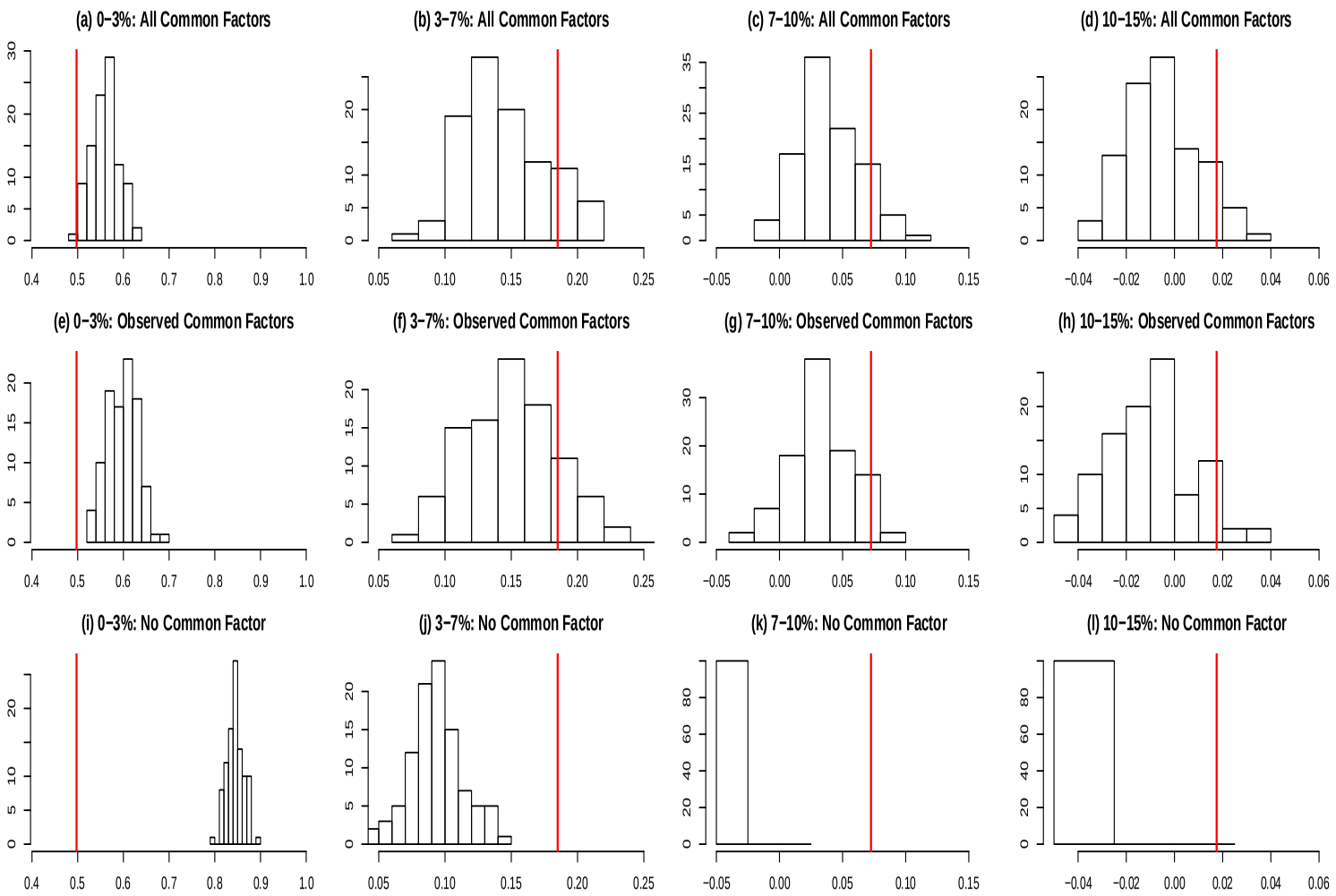

Figure B.14: **Histograms of the Posterior Samples of CDX Tranche Prices on September 23, 2009 When We Use Rolling Data Window under the Structural Model.** We estimate CDX tranche prices on September 23 using the data range from April 22 to September 22. Histograms (a)-(d) depict the sampled posterior distributions of each tranche price ((from the left) 0-3%, 3-7%, 7-10% and 10-15%), when we consider all common factors. Histograms (e)-(h) depict the sampled posterior distributions of each tranche price, when we consider only observed common factors. Histograms (i)-(l) depict the sampled posterior distributions of each tranche price, when we do not use any of the common factors (independent model).

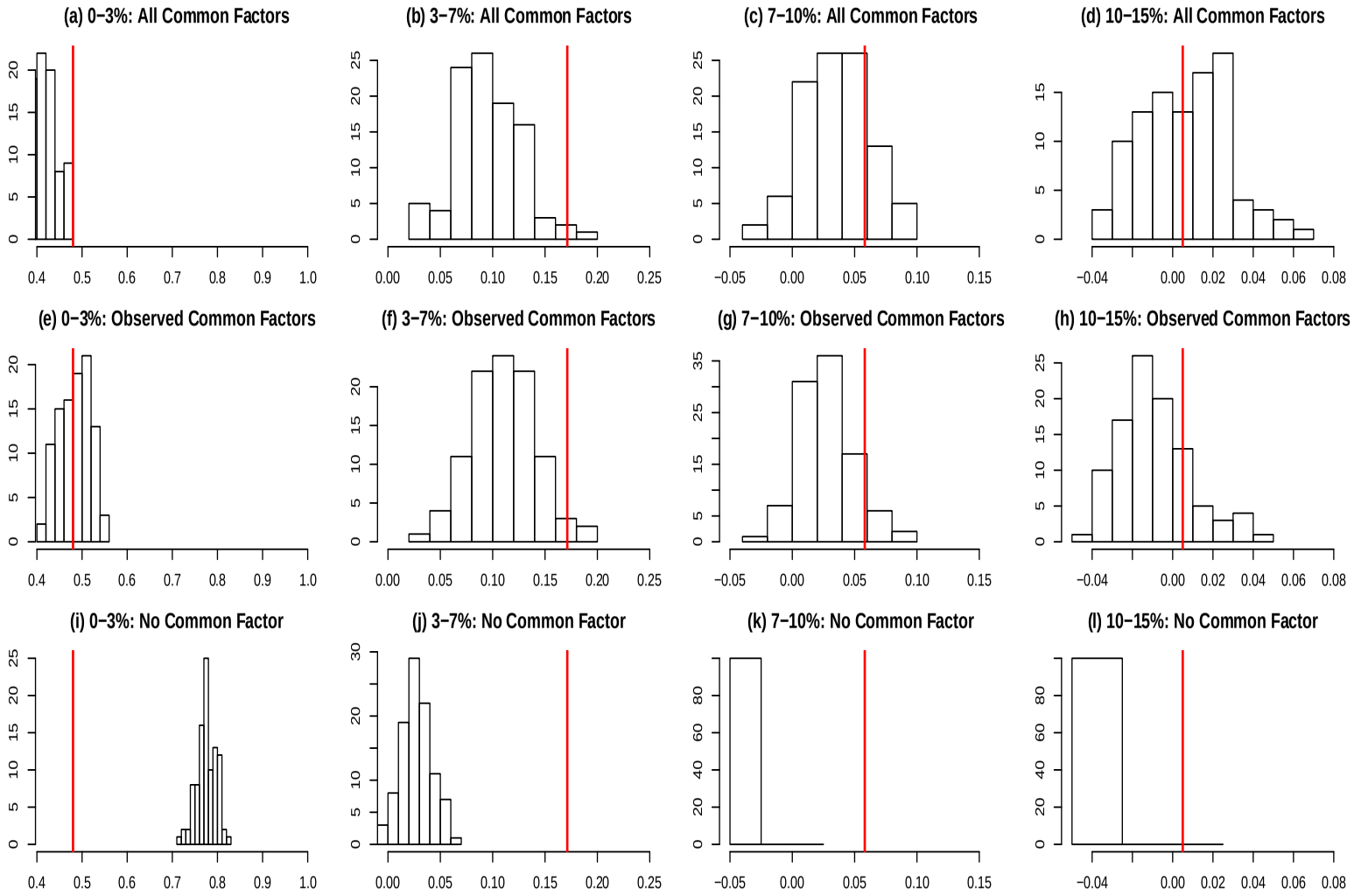

Figure B.15: **Histograms of the Posterior Samples of CDX Tranche Prices on December 31, 2009 When We Use Rolling Data Window under the Structural Model:** We estimate CDX tranche prices on December 31 using the data range from September 30 to December 30. Histograms (a)-(d) depict the sampled posterior distributions of each tranche price ((from the left) 0-3%, 3-7%, 7-10% and 10-15%), when we consider all common factors. Histograms (e)-(h) depict the sampled posterior distributions of each tranche price, when we consider only observed common factors. Histograms (i)-(l) depict the sampled posterior distributions of each tranche price, when we do not use any of the common factors (independent model).

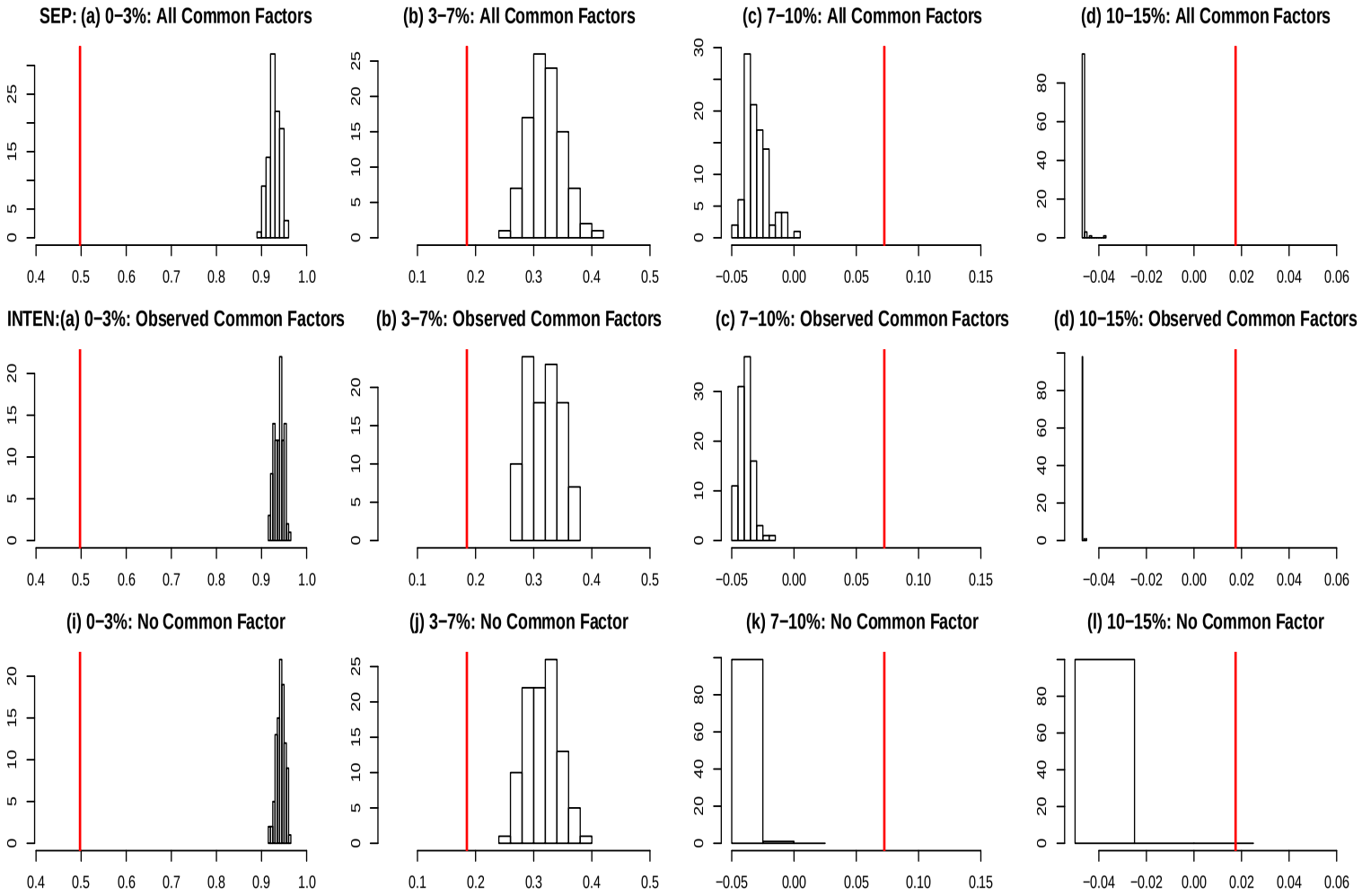

Figure B.16: **Histograms of the Posterior Samples of CDX Tranche Prices on September 23, 2009 When We Use Rolling Data Window under the Intensity Model:** We estimate CDX tranche prices on September 23 using the data range from April 22 to September 22. Histograms (a)-(d) depict the sampled posterior distributions of each tranche price ((from the left) 0-3%, 3-7%, 7-10% and 10-15%), when we consider all common factors. Histograms (e)-(h) depict the sampled posterior distributions of each tranche price, when we consider only observed common factors. Histograms (i)-(l) depict the sampled posterior distributions of each tranche price, when we do not use any of the common factors (independent model).

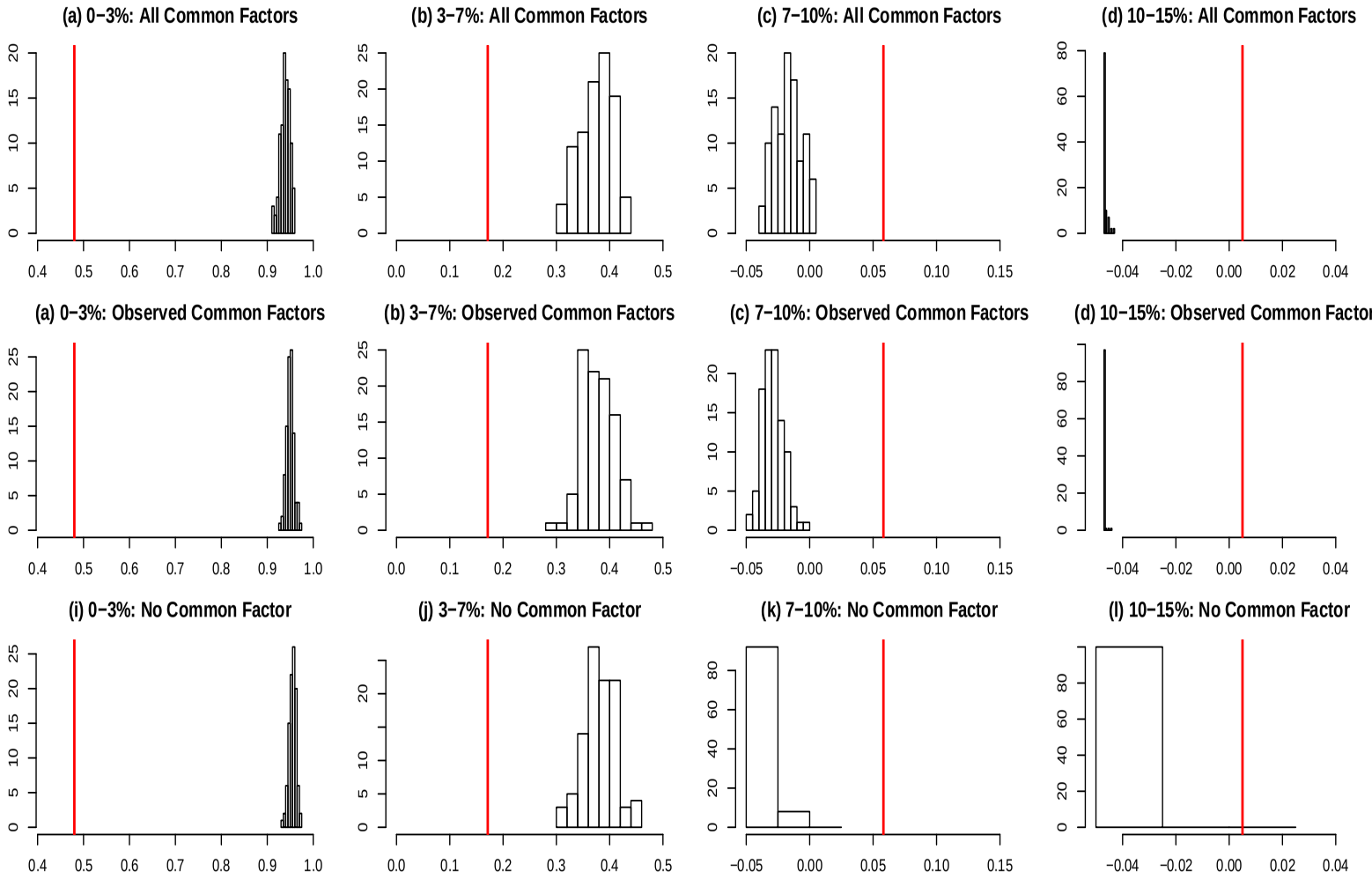

Figure B.17: **Histograms of the Posterior Samples of CDX Tranche Prices on December 31, 2009 When We Use Rolling Data Window under the Intensity Model:** We estimate CDX tranche prices on December 31 using the data range from September 30 to December 30. Histograms (a)-(d) depict the sampled posterior distributions of each tranche price ((from the left) 0-3%, 3-7%, 7-10% and 10-15%), when we consider all common factors. Histograms (e)-(h) depict the sampled posterior distributions of each tranche price, when we consider only observed common factors. Histograms (i)-(l) depict the sampled posterior distributions of each tranche price, when we do not use any of the common factors (independent model).

BIBLIOGRAPHY

Ait-Sahalia, Y., Mykland P., and L. Zhang, 2005a, How Often to Sample a Continuous-Time Process in the Presence of Market Microstructure Noise, *Review of Financial Studies* 18, 351–416.

Ait-Sahalia, Y., Mykland P., and L. Zhang, 2005b, A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data, *Journal of the American Statistical Association* 100, 1394–1411.

Ait-Sahalia, Y., and J. Yu, 2009, High Frequency Market Microstructure Noise Estimates and Liquidity Measures, *The Annals of Applied Statistics* 3, 422–457.

Arora, N., J.R. Bohn, and F. Zhu, 2005, Reduced Form vs. Structural Models of Credit Risk: A Case Study of Three Models, *White Paper, Moody's KMV.*

Bandi, F., and J. Russell, 2006, Separating Microstructure Noise from Volatility, *Journal of Financial Economics* 79, 655–692.

Bielecki, T.R., M. Jeanblanc, and M. Rutkowski, 2011, Hedging of a Credit Default Swaption in the CIR Default Intensity Model, *Finance and Stochastics* 15.

Black, F., and J.C. Cox, 1976, Valuing Corporate Securities: Some Effects of Bond Indenture Provisions, *Journal of Finance* 31, 351–67.

Black, F., and M. Scholes, 1973, The Pricing of Options and Corporate Liabilities, *The Journal of Political Economy* 81, 637–654.

Brigo, D., and A. Alfonsi, 2005, Credit Default Swaps Calibration and Option Pricing with the SSRD Stochastic Intensity and Interest-Rate Model, *Finance Stoch.* 9, 29–42.

Brooks, S.P., and A. Gelman, 1998, General Methods for Monitoring Convergence of Iterative Simulations., *Journal of Computational and Graphical Statistics* 7, 434–455.

Campbell, J.Y., and M. Yogo, 2006, Efficient Tests of Stock Return Predictability, *Journal of Financial Economics* 81, 27–60.

Collin-Dufresne, P., and J. Helwege, 2003, Is Credit Event Risk Priced? Modeling Contagion via the Updating of Beliefs, *Working Paper*.

Conrad, J., and G. Kaul, 1989, Mean Reversion in Short-Horizon Expected Returns, *Review of Financial Studies* 2, 225–240.

Das, S., and P. Hanouna, 2009, Implied Recovery, *Journal of Economic Dynamics & Control* 133, 1837–1857.

Daveri, F., 2009, Italy, Before and After Lehman Brothers, *VoxEU.org*.

De Servigny, A, and O. Renault, 2002, Default Correlation: Empirical Evidence, *Standard& Poors Risk Solutions*.

Duan, J.C., 1994, Maximum Likelihood Estimation Using Price Data of the Derivative Contract, *Mathematical Finance* 4, 155–167.

Duan, J.C., 2000, Correction: Maximum Likelihood Estimation Using Price Data of the Derivative Contract, *Mathematical Finance* 10, 461–462.

Duan, J.C., and A. Fulop, 2009, Estimating the Structural Credit Risk Model When Equity Prices Are Contaminated by Trading Noises, *Journal of Econometrics* 150, 288–296.

Duffee, G.R., 1999, Estimating the Price of Default Risk, *The Review of Financial Studies* 12, 197–226.

Duffie, D., A. Eckner, G. Horel, and L. Saita, 2009, Frailty Correlated Default, *Journal of Finance* 64, 2089–2123.

Duffie, D, and N. Garleanu, 2001, Risk and Valuation of Collateralized Debt Obligations, *Financial Analysts Journal* 57, 41–59.

Duffie, D., Saita L., and K. Wang, 2007, Multi-Period Corporate Default Prediction with Stochastic Covariates, *Journal of Financial Economics* 83, 735–665.

Fama, E.F., and K. French, 1988, Permanent and Temporary Components of Stock Prices, *Journal of Political Economy* 96, 246–273.

Fama, E.F., and K.R. French, 1996a, Multifactor Explanations of Asset Pricing Anomalies, *The Journal of Finance* 51, 55–84.

Fama, E.F., and K.R. French, 1996b, Multifactor Portfolio Efficiency and Multifactor Asset Pricing, *Journal of Financial and Quantitative Analysis* 31, 441–465.

Feldhutter, P, and M.S. Nielsen, 2012, Systematic and Idiosyncratic Default Risk in Synthetic Credit Markets, *Journal of Financial Econometrics* pp. 292–324.

Fleming, M.J., 2003, Measuring Treasury Market Liquidity, *FRBNY Economic Policy Review* pp. 83–108.

Friewald, N, 2009, Estimating Asset Correlation from CDS Spreads: An Analysis of the Hedge Effectiveness of FTD Basket, *Working Paper*.

Gaunt, C., and P. Gray, 2003, Short-Term Autocorrelation in Australian Equities, *Australian Journal of Management* 28, 97–117.

Gelman, A., and J. Hill, 2007, *Data Analysis Using Regression and Multilevel/Herachical Models*. (Cambridge Boca Ranton, London, New York and Washington, D.C.).

Gelman, A., and D.B. Rubin, 1992, Inference From Iterative Simulation using Multiple Sequences, *Statistical Science* 7, 457–511.

Gordon, N. J., D.J. Salmond, and A.F.M. Smith, 1993, A Novel Approach to Non-linear and Non-Gaussian Bayesian State Estimation, *IEE-Proceedings* F140, 107–113.

Gordy, M., 2003, A Risk-Factor Model Foundation for Ratings-Based Bank Capital Rules, *Journal of Financial Intermediation* 12, 199–232.

Hull, J., I. Nelken, and A. White, 2004, Merton's Model, Credit Risk and Volatility Skews, *Journal of Credit Risk* 1, 3–28.

Hull, J., M. Predescu, and A. White, 2010, The Valuation of Correlation-Dependent Credit Derivatives Using a Structural Model, *Journal of Credit Risk* 6.

Hull, J., and A. White, 1987, The Pricing of Options on Assets with Stochastic Volatilities, *The Journal of Finance* 2, 281–300.

Hull, J., and A. White, 2000, Valuing Credit Default Swaps I: No Counterparty Default Risk, *Journal of Derivatives* 8, 29–40.

Hull, J., and A. White, 2004, Valuation of a CDO and an n^{th} to Default CDS without Monte Carlo Simulation, *Journal of Derivatives* 12, 8–23.

Ito, K., 1951, On Stochastic Differential Equations, *Memoirs of the American Mathematical Society* 4, 1–51.

Jegadeesh, N., 1990, Evidence of the Predictable Behavior of Security Returns, *Journal of Finance* 45, 881–898.

Johannes, M. S., N. G. Polson, and J. R. Stroud, 2007, Optimal Filtering of Jump Diffusions: Extracting Latent States from Asset Prices, *Working Paper*.

Kou, S. C., S.X. Xie, and J.S. Liu, 2005, Bayesian Analysis of Single-Molecule Experimental Data, *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 54, 469–506.

K.P., Michael, and N. Shephard, 1999, Filtering via Simulation: Auxiliary Particle Filters, *Journal of the American Statistical Association* 94, 590–599.

Lando, D., 2004, *Credit Risk Modeling: Theory and Applications*. (Princeton University Press Boston, MA).

Lehmann, B., 1990, Fads, Martingales, and Market Efficiency, *Quarterly Journal of Economics* 105, 1–28.

Liang, J., J. Ma, T. Wang, and Q. Ji, 2011, Valuation of Portfolio Credit Derivatives with Default Intensities Using the Vasicek Model, *Asia-Pacific Finan Markets* 18, 33–54.

Liu, J.S., 1996, Metropolized Independent Sampling with Comparisons to Rejection Sampling and Important Sampling, *Statistics and Computing* 6.

Liu, J.S., and C. Sabatti, 2000, Generalized Gibbs Sampler and Multigrid Monte Carlo for Bayesian Computation, *Biometrika* 87, 353–369.

Lo, A.W., and A.C. MacKinlay, 1988, Stock Market Prices Do Not Follow Random Walks: Evidence From a Simple Specification Test, *Review of Financial Studies* 1, 41–66.

Longstaff, F.A., and A. Rajan, 2008, An Empirical Analysis of the Pricing of Collateralized Debt Obligations, *The Journal of Finance* 63, 529–563.

Lucas, D.J., 1995, Default Correlation and Credit Analysis, *Journal of Fixed Income* March, 76–87.

Maitland, M., and D. Blitzer, 2002, A GICS Overview for Standard & Poor's U.S. Indices, *Standard & Poor's Quantitative Services*.

Mcculloch, C. E., and S. R. Searle, 2001, *Generalized, Linear and Mixed Models*. (John Wiley & Sons.).

Merton, R.C., 1974, On the Pricing of Corporate Debt: The Risk Structure of Interest Rates, *Journal of Finance* 29, 449–70.

Mortensen, A., 2006, Semi-Analytical Valuation of Basket Credit Derivatives in Intensity-Based Models, *Journal of Derivatives* pp. 8–26.

Ni, S.X., J. Pan, and A.M. Poteshman, 2008, Volatility Information Trading in the Option Market, *Journal of Finance* 63, 1059–1091.

Pan, J., and K.J. Singleton, 2008, Default and Recovery Implicit in the Term Structure of Sovereign CDS Spreads, *The Journal of Finance* 63, 2345–2384.

Poterba, J., and L. Summers, 1988, Mean Reversion in Stock Returns: Evidence and Implications, *Journal of Financial Economics* 22, 27–60.

Rubin, D. B., 1987, Comment on 'The Calculation of Posterior Distributions by Data Augmentation' by M. A. Tanner and W. H. Wong, *Journal of the American Statistical Association* 82, 543–546.

Shumway, T., 2001, Forecasting Bankruptcy More Accurately: A Simple Hazard Model, *Journal of Business* 74, 101–124.

S.L., Jun, 2001, *Monte Carlo Strategies in Scientific Computing*. (Springer Cambridge, MA).

Smith, R., and I. Walter, 2006, Four Years After Enron: Assessing the Financial-Market Regulatory Cleanup, *The Independent Review* 11, 53–66.

Torresetti, R., D. Brigo, and A. Pallavicini, 2007, Implied Expected Tranched Loss Surface from CDO Data, *Credit Models - Banca IMI Corso Matteotti 6 - 20121 Milano, Italy*.

Turnbull, S.M., and J. Yang, 2008, Default Dependence: The Equity Default Relationship, *EFA 2008 Athens Meetings Paper*.

Vasicek, O.A., 1991, Limiting Loan Loss Probability Distribution, *KMV Corporation*.

Vassalou, M., and Y. Xing, 2004, Default Risk in Equity Returns, *Journal of Finance* 59, 831–868.

Wei, G.C., and M.A. Tanner, 1990, A Monte Carlo Implementation of the EM Algorithm and the Poor Man's Data Augmentation Algorithms, *Journal of the American Statistical Association* 85, 699–704.

Zhou, C., 2001, The Term Structure of Credit Spreads with Jump Risk, *Journal of Banking & Finance* 25, 2015–2040.

Zipunnikov, V.V., and J.G. Booth, 2006, Monte Carlo EM for Generalized Linear Mixed Models Using Randomized Spherical Radial Integration, *Working Paper, Cornell University*.